INDIVIDUAL AND COMMON INFORMATION: MODEL-FREE EVIDENCE FROM PROBABILITY FORECASTS

YIZHOU KUANG AND KRISTOFFER NIMARK

Abstract. We decompose cross-sections of belief revisions from the SPF into common and individual components. The common component is modeled *as if* caused by the single (hypothetical) signal that, if observed by all forecasters, can explain the maximum amount of belief revisions across forecasters. The residual belief revisions not accounted for by the common component are modeled as if caused by forecaster specific signals. On average, individual information accounts for more of the observed belief revisions than common information, but with large cross-sectional heterogeneity. Inflation volatility, perceived stock market volatility and a high risk of recession are all associated with increased informativeness of signals. The cyclicality of informativeness is most pronounced in individual signals except during the COVID pandemic when the informativeness of common signals increased sharply.

1. Introduction

Decisions taken under uncertainty can be improved upon by having more information and how, when, and for what purposes economic agents acquire information is the subject of a large and active theoretical literature. From this literature, we know that information that is common to many agents is more likely to affect economic aggregates and that whether information is private or public is of particular importance in strategic environments.[1] In spite of these important distinctions, there is very little empirical work studying the relative importance of individual and common information acquisition outside highly structural models. In this paper we aim to make two contributions to remedy this short-coming. First, we propose a method that allows us to extract individual and common information from repeated fixed-event probability forecasts. Second, we demonstrate how the method can be used to ask and answer new questions about the empirical properties of individual and common information.

The proposed procedure can be applied whenever we can observe a panel of probability forecasts about a fixed event. It contains two steps. In the first step, we find the single signal that, if observed by all forecasters, can explain the maximum amount of the cross-section of belief revisions. This signal is defined as the common signal. In the second step,

Date: April 11, 2025. The authors thank Etienne Briand, Ryan Chahrour, Alessandro Ferrari, Zhonghao Fu, Nikolay Iskrev, Oleksiy Kryvtsov, Ricardo Reis, Alireza Tahbaz-Salehi, Andreas Uthemann, Zebang Xu and seminar participants at Cornell University, the University of Zurich, Bank of Canada, the First Paris Conference on the Macroeconomics of Expectations (Banque de France-PSE-CEPR), the University of New South Wales and UQAM for useful comments and suggestions.

Kuang: yizhou.kuang@manchester.ac.uk, https://yizhoukuang.com; *Nimark*: pkn8@cornell.edu, www.kris-nimark.net.

[1]See for instance Morris and Shin (2002) or Hellwig and Veldkamp (2009).

we invert Bayes rule to extract the implied individual signal for each forecaster so that when combined with the common signal, the two types of signals completely account for the observed cross-section of belief revisions. The method imposes relatively weak assumptions, namely that forecasters update their beliefs using Bayes rule. We impose neither that the observed beliefs are rational, follow a particular parametric functional form, nor that the information structure is stable over time.[2]

It is important to note here that we extract the common component of belief revisions *as if* it is completely explained by a single commonly observed signal. In practice, the common component in belief revisions may be influenced by multiple common signals observed within the period, or as we will show later, by the presence of a common component in privately observed signals. In spirit, the method is thus akin to extracting the first-principal component from a cross-section of belief revisions and then labeling the remaining belief revisions as idiosyncratic. The latter are thus modeled *as if* it is driven by a single individual signal for each forecaster. It is only under special circumstances that the extracted signals correspond exactly to a theoretical model object or variable. Similarly to principal components, the extracted signals are thus defined by their statistical properties and not by an underlying economic model structure.

In strategic settings, it is also important to distinguish between common information and public information, where the latter is not only known by all agents but also common knowledge. Our approach does not allow us to distinguish between common and public information. Since all public information is also common information, but not vice versa, we use the weaker terms *common information* and *common signals* throughout.

When we apply the proposed method to probability forecasts from the *Survey of Professional Forecasters* (SPF), we find that (i) individual information is on average more informative than common information, and account for more of the observed belief revisions, (ii) there is a large cross-sectional heterogeneity in signal informativeness, and the fraction of forecasters that observe individual signals that are more informative than the common signal ranges from 0.1 - 0.9, depending on variable and measure of informativeness, (iii) volatile inflation is associated with more informative signals (iv) the precision of signals tends to increase when the probability of a recession is high, or when the perceived volatility of stock prices is high.

The SPF measures forecasters' subjective beliefs, and the proposed method thus only allows us to measure how the perceived informativeness of individual and common sources of information vary over time. However, as shown in Bassetti, Casarin and Del Negro (2022), at short horizons, more precise beliefs correspond to a higher actual precision of forecasts. Whether the distinction that we measure forecasters' subjective uncertainty matters or not depends on the purpose of the analysis. For instance, if we are interested in whether forecasters form beliefs rationally, then it certainly matters whether forecasts are more accurate when they perceive them to be so. However, for understanding decisions taken under uncertainty, it is arguably the subjective uncertainty that is more relevant.

[2]Giacomini, Skreta and Turen (2020) provides empirical support for the assumption that forecasters use Bayes rule to update their beliefs.

To understand how the method works, it is helpful to first delve a little bit into the structure of the SPF. The administrators of the survey collect both point and probability forecasts and for this study we make use of the latter. The SPF asks respondents to assign probabilities to different ranges ("bins") of outcomes for GDP growth, the GDP deflator, Personal Consumption Expenditure (PCE) inflation, Consumer Price Index (CPI) inflation and unemployment. The bins are pre-specified by the SPF and occasionally redefined due to changes in the long term means and variances of the variables. Both point and probability forecasts are collected every quarter. However, unlike the point forecasts, the probability forecasts are only elicited about calendar year outcomes, i.e. they are fixed-event rather than fixed-horizon forecasts.

The fixed-event nature of the probability forecasts allows us to study how forecasters revise their beliefs about a given event over time. In particular, since calendar year forecasts are collected every quarter and for multiple calendar years at each survey wave, we can observe how the cross-section of beliefs about a given calendar year changes quarter-to-quarter. For instance, survey respondents are asked to provide a probability forecasts for CPI inflation for the calendar year 2015 every quarter from 2012:Q1 to 2015:Q4. Hence, we have 16 cross-sections of probability forecasts about CPI inflation for the calendar year 2015 and 15 observed cross-sections of revisions to these beliefs.

We use this structure to estimate the relative importance of common and individual information in the observed belief revisions. The basic idea is the following. For a given change in a forecaster's probability forecast, we can invert Bayes rule to back out a signal that would justify the change in his beliefs from $t-1$ to t. If done individually for each forecaster, we would end up with one signal for each forecaster at each point in time. However, we want to separate out the component of each forecaster's belief revision that is due to common information from the component that is due to individual information. To do so, we ask *What is the single signal that, if observed by all forecasters, can explain the most of the belief revisions of all the forecasters?* We call this signal the *common signal*. To make the procedure operational, we find the signal that minimizes the Kullback-Leibler divergence between the observed cross-section of beliefs in period t and the hypothetical cross-section of beliefs forecasters would have had, had they updated their prior based only on the common signal.

The extracted common signal will not by itself in general be enough to completely account for how every forecaster updates his or her beliefs. However, we can back out an implied individual signal that when combined with the observed prior and the common signal, completely account for a given forecaster's observed belief revision.

To study the importance of individual and common information and how it varies over time, we propose three measures of signal informativeness. The *belief update measure* captures how large a revision of an agent's belief a signal leads to. While a natural measure of a signal's importance, the belief update measure depends on the prior of the agent and not only on the properties of the signal. The *negative entropy measure* is independent of forecasters' priors and measures how much a signal reduces the entropy over possible outcomes from a starting point of maximum entropy. The belief update and entropy measures are independent of the numerical values associated with different outcomes and are thus not suited to measure the precision of a signal. However, the *precision measure* of a signal, computed as the inverse of the variance of the hypothetical posterior implied by combining the signal with a uniform prior, allows us to evaluate the perceived precision of signals. Importantly, the three measures do not necessarily comove positively. For instance, a signal that suggests that tail events are more likely may cause a large increase in a the belief update measure, but a large decrease in the precision measure. The procedure also allow for that relatively precise individual and common signals may increase posterior uncertainty if they disagree about the conditional mean of an outcome. The precision measure of the signal informativeness of the two types of signals are thus not simply a proxy for changes in posterior uncertainty. This also distinguishes the method from studies that use a fully parametric linear-Gaussian statistical model to make inference about agents beliefs and signals, i.e. Barillas and Nimark (2017, 2019) and Nimark (2014). In such models, all signal realizations about a fixed event weakly decreases posterior uncertainty, even when signals individually imply very different conditional means.

We document several empirical regularities about the extracted signals. First, individual signals are for most agents more informative than the public signals, in the sense that the individual signals account for more of the observed belief revisions and are perceived to be more precise. This is true in spite of the fact that the procedure used to extract the common signal maximizes the importance of the common signals. While the average informativeness of individual signals is higher than that of the common signal, there is substantial cross-sectional heterogeneity. The fraction of forecasters that observe individual signals that are more informative than the common signal ranges from 0.1 - 0.9, depending on variable and measure of informativeness. For all variables and measures, except the precision measure for signals about unemployment and GDP growth, the common signals are less informative than the individual signals for a majority of forecasters. Again, this is so in spite of the procedure maximizing the informativeness of the common signal.

For some macro variables, the outcomes of the underlying variable that is being forecast covary with the informativeness of the signals. For example, volatile inflation tends to be associated with both individual and common signals becoming more precise. This is consistent with theories of agents rationally choosing how much attention to pay to inflation, as analyzed in Pfauti (2023) and discussed by Federal Reserve chair Jerome Powell in a recent speech.[3] High levels of unemployment tend to be associated with more precise individual and common signals. Interestingly, when unemployment is increasing, both individual and common signals about unemployment tend to be less precise. One possible interpretation of this result is that in a recession, when unemployment increases rapidly, there is an increased uncertainty about how high unemployment will go before it peaks.

Signal informativeness are almost uniformly positively correlated with the Philadelphia Fed's *Anxious Index,* which describes forecasters' subjective probability of a recession. Perhaps unsurprisingly, this is most significantly so for the common signals about unemployment and GDP growth. A similar pattern holds for the correlations between the measures of signal informativeness and the *VIX* index from the Chicago Board Options Exchange, with the addition that the correlation between informativeness of the individual signals and CPI

[3]In his August 26, 2022 speech *on Monetary Policy and Price Stability,* Federal Reserve chair Jerome Powell said "*When inflation is persistently high, households and businesses must pay close attention and incorporate inflation into their economic decisions. When inflation is low and stable, they are freer to focus their attention elsewhere.*".

inflation is stronger for the VIX than for the *Anxious Index*. These findings suggest that in times of either increased probability of a recession, or when there is a high level of perceived uncertainty about the stock market, the incentives to acquire information by the survey participants may be particularly strong. The latter is consistent with the mechanisms explored in Song and Stern (2020), Flynn and Sastry (2022) and Chiang (2022), who all argue that firms have a stronger incentive to acquire information in bad times. Flynn and Sastry (2022) further argue that this fact can explain why we observe asymmetric business cycles with state-dependent dynamics. The evidence here is also consistent with the empirical findings in Song and Stern (2020) and Flynn and Sastry (2022) who both use a text-based approach to measure firms' attention to macroeconomic variables and find it to be counter-cyclical.

There exists a large empirical literature studying the Survey of Professional Forecasters. One strand of this literature has focused on the accuracy of the forecasts, and in particular their accuracy relative to alternative econometric forecasting models, e.g. Zarnowitz (1979), Zarnowitz and Braun (1993), Diebold, Tay, and Wallis (1997), Clements (2006, 2018), Engelberg, Manski and Williams (2009) and Kenny, Kostka and Masera (2014). A second strand has studied how to best combine individual survey forecasts to increase forecast accuracy, e.g. Bonham and Cohen (2001) and Genre, Kenny, Meyler and Timmermann (2013). A third strand has focused on testing theories of expectations formation, including the rational expectations hypothesis, e.g. Zarnowitz (1985), Keane and Runkle (1990), Bonham and Dacy (1991), Laster, Bennett and Geoum (1999) and Coibion and Gorodnichenko (2012, 2015).

Most of the literature using or studying the SPF focuses on the point forecasts, which are available for a larger set of macro variables. Some exceptions that do make use of the probability forecasts include Diebold, Tay and Wallis (1997), Clements (2006), Kenny, Kostka and Masera (2014), Rossi, Sekhposyan, and Soupre (2019), Clements (2018), Ganics, Rossi, and Sekhposyan (2020), Bassetti, Casarin, and Del Negro (2022) and Clements, Rich and Tracy (2025). These studies mostly focus on the the accuracy of the forecasts. Rossi *et al* (2019) decompose SPF probability forecasts into "Knightian uncertainty" and "risk" components, while Ganics *et al* (2020) propose a method to construct fixed-horizon probability forecasts from fixed-event forecasts.[4]

Manzan (2021) and Clements, Rich and Tracy (2025) use the SPF probability forecasts to test whether forecasters' uncertainty decreases with the forecast horizon, which should be the case if the forecasters are Bayesian and the true underlying information structure is linear-Gaussian. Lahiri and Sheng (2010) derive conditions for the relationship between uncertainty and cross-sectional dispersion of forecasts that hold in a linear-Gaussian setting. By not imposing a particular parametric form for the underlying information structure, we can separately study disagreement and uncertainty. Importantly, our framework allows for some signals to be both informative and to increase uncertainty, something that is ruled out in a linear-Gaussian setting where all informative signals must also reduce uncertainty.

[4] Rossi et al (2016) assume that there exists an objective "correct" probability distribution over future outcomes and interpret dispersion of distributions across forecasters as evidence of Knightian uncertainty. Here, we interpret dispersion of beliefs across forecasters as arising from differences in information sets.

The next section describes the structure of the probability forecasts in the SPF. Section 3 proposes a procedure to extract individual and common signals from a cross-section of belief revisions. Section 4 characterizes the estimated common and individual information components and shows how they map into alternative information structures. Section 5 proposes three measures of signal informativeness and Section 6 presents the empirical results. Section 7 concludes.

2. The structure of the SPF probability forecast data

The Survey of Professional Forecasters (SPF) contains quarterly forecasts from practitioners in industry, Wall Street, commercial banks and academic research centers about key macroeconomic variables. Since 1990 it has been administered by the Federal Reserve Bank of Philadelphia who took over the survey from the American Statistical Association and the National Bureau of Economic Research. All participants produce forecasts as part of their current jobs. The respondents are anonymous to users of the survey, but individual forecasters can be tracked over time through an id number.[5] The SPF collects both point forecasts and probability forecasts. The point forecasts have been used widely to study properties of expectations formation and as a benchmark for evaluating statistical forecasting models and procedures. The probability forecasts, like the point forecasts, are collected every quarter. However, the SPF only elicit probability forecasts for a subset of the macro variables that they elicit point forecasts for, and all probability forecasts are fixed-event rather than fixed-horizon forecasts.

The SPF currently collects probability forecasts about the GDP growth rate, the GDP deflator inflation, CPI inflation, PCE inflation and the unemployment rate. The longest sample is available for the GDP growth rate and the GDP deflator inflation, starting in 1968:Q4. However, until 1981:Q3, respondents were only asked to provide probability forecasts for the current calendar year. Since 1981:Q4 they have been asked to also provide probability forecasts for the next calendar year, and since 2009:Q2 they have been asked to forecast the next three calendar years in addition to the current one. Probability forecasts for CPI inflation and PCE inflation have been included in the survey since 2007:Q2 and probability forecasts for unemployment were added in 2009:Q2.

The probability forecasts for the variables added since 2007 include forecasts for the current year, as well as forecasts for the next three calendar years. Respondents are asked to assign probabilities to ranges ("bins") of different outcomes, where the intervals defining each bin is predefined by the administrators of the survey. The definitions of the bins have occasionally been changed to reflect that the high-probability ranges of the macro variables have changed. The survey responses and the bin definitions are illustrated in Figure 2.1 where we have plotted the average probability forecasts for the next calendar year outcome of the five macroeconomic variables. The x-axis denotes the quarter when the forecast was made. The y-axis denotes the outcomes and horizontal dotted lines indicate bin boundaries. A vertical line indicates a date when a redefinition of the bins occurred. These redefinitions have been motivated by either a persistent change in the mean or variance of the probability forecasts

[5] For a detailed description of the survey and how it has changed over time, see Croushore (1993) and Croushore and Stark (2019).

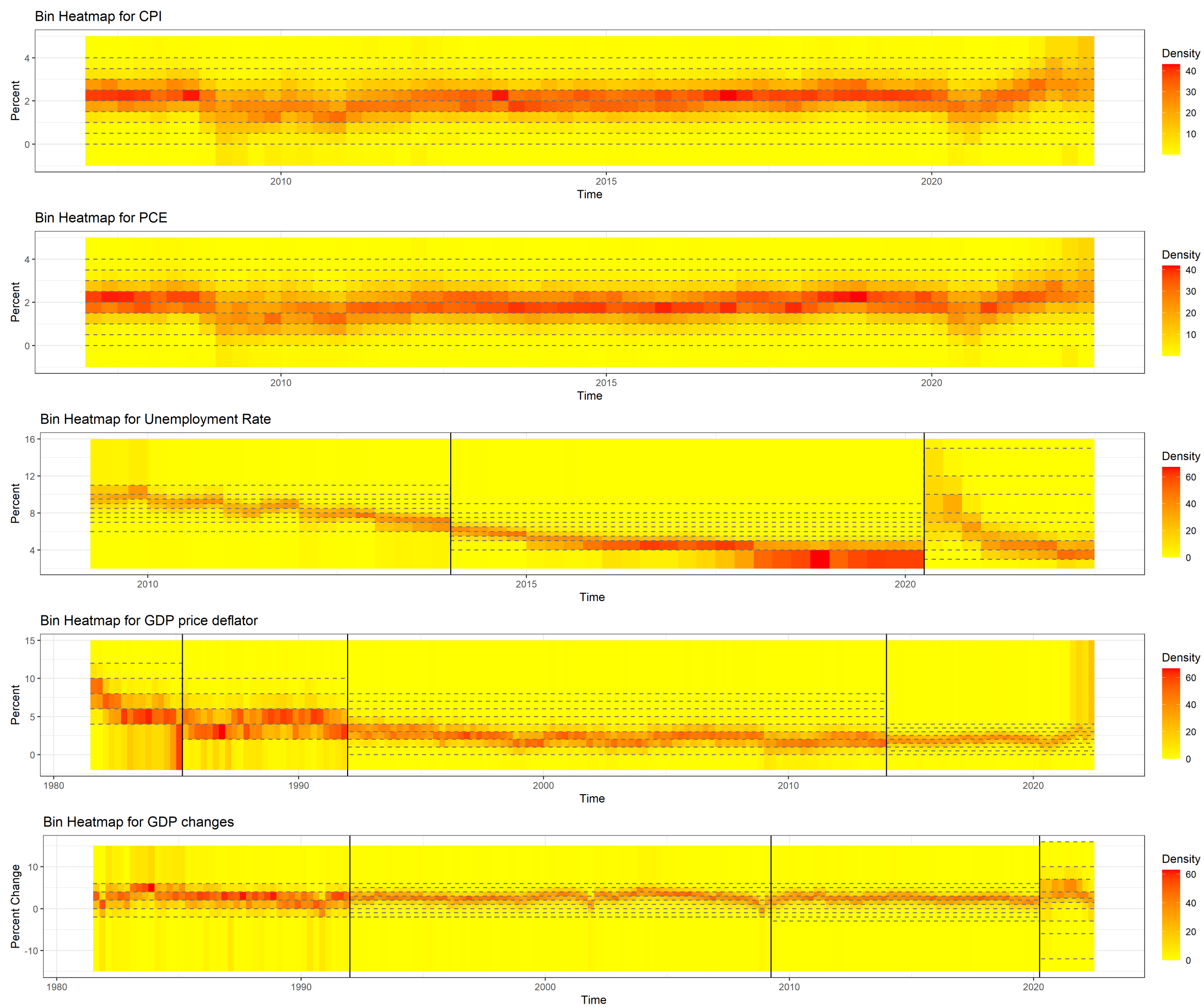

FIGURE 2.1. Average density forecast for CPI, PCE, unemployment, GDP price deflator and GDP growth. Dashed horizontal lines indicate bin boundaries, vertical solid lines indicate date of change for bin definitions.

or, in the case of the redefinitions of the bins for unemployment and GDP growth during 2020, to address an abrupt change in the plausible range of outcomes. When extracting a signal based on updates across a bin change, we convert each distribution into the coarsest common bin definition before extracting the signal.

Respondents are asked to repeatedly, i.e. over several consecutive quarters, forecast a given calendar year outcome. The fixed-event nature of the probability forecasts implies that we can observe how a forecaster's beliefs about a given calendar year outcome evolves over time. This is illustrated in Figure 2.2 where we have plotted how the probability forecast of Forecaster #570 about CPI inflation in 2021 changed quarter-to-quarter in 2020. When COVID struck in 2020:Q2, the forecaster shifted the distribution to the right, increasing the probability of high inflation outcomes and in subsequent quarter, he or she continued to

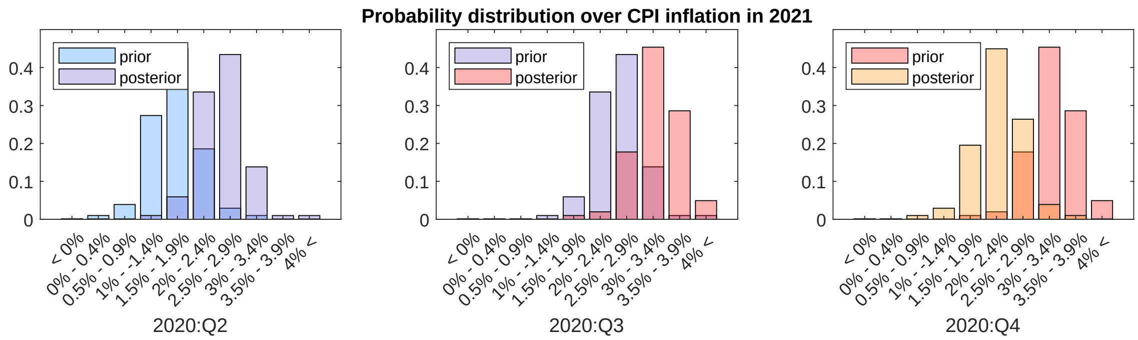

FIGURE 2.2. Illustration of how the beliefs of Forecaster #570 about inflation in 2021 evolved over 2020. A given color indicates beliefs reported in a given quarter, sequenced as blue→purple→red→yellow. Priors are equal to the posteriors from the previous quarter.

shift even more probability mass to high inflation outcomes. In the last quarter, the revisions changed directions and the forecaster then increased the probability of more moderate inflation outcomes.

3. Extracting common and individual information from probability forecasts

The Survey of Professional Forecasters (SPF) allows us to observe how a cross-section of individual forecasters' beliefs about a single future event change over time.[6] The basic idea is to find the single signal that, if observed by everyone, explains "the most" of the observed belief revisions. Then, individual signals are defined to explain any residual revisions not accounted for by the common signal. Below, we describe how to implement this idea operationally, but first we offer a note on interpretation, notation, and terminology.

As noted in the introduction, the common signal is extracted *as if* the common component in the belief revision was explained by a single signal. Likewise, the individual signal is extracted *as if* the idiosyncratic belief revision of an individual agent was caused by a single signal. In practice, both types of signals capture beliefs revisions most likely driven by information from multiple sources.

We index forecasters by $j \in 1, 2, ..., J$ and time by $t \in 1, 2, ..., T$. We use x to denote a generic macroeconomic outcome which can take values in $n \in 1, 2, ..., N$ different intervals (or bins) in X. The probability forecast of forecaster j in period t is denoted $p(\boldsymbol{x} \mid \Omega_t^j)$ where $\boldsymbol{x} \equiv (x_1, x_2, ..., x_N)$ is the vector of possible outcomes and Ω_t^j denotes the information set

[6]In practice, however, many surveys report forecasts in a fixed-horizon format rather than a fixed-event format. To apply the method to such data, one could simply transform a panel of fixed-horizon probability forecasts into a (shorter) panel of fixed-event forecasts by "lining up" the forecasts that target the same calendar date outcome. Effectively, this could be achieved by using the method proposed by Ganics, Rossi, and Sekhposyan (2024) to convert fixed-event into fixed-horizon distributions in reverse.

available to forecaster j in period t. A signal structure is a probability distribution $p(S \mid X)$ that associates a probability for each possible signal $s \in S$ with each possible outcome $x \in X$.

It is natural to formulate the discussion below in terms of prior and posterior distributions, where the posterior is obtained by combining the new information in the signal with the information in the prior. However, it is worth remembering that given the fixed-event nature of the forecasts, the prior in period t is simply the posterior inherited from period $t-1$.

3.1. **Bayes' rule, belief updates and realized signals.** We do not observe neither forecasters' signal structures, nor the realized signals directly, but observing both the prior and the posterior allow us to infer the implied relative probability of observing the realized signal in different states $x \in X$.

To understand what we can learn about the properties of realized signals from observing beliefs revisions, it is helpful to start with the mechanics of Bayesian updating. Denote the prior of forecaster j as Ω^j_{t-1} and the signal observed in period t as s_t.[7] Bayes rule then tells us that the posterior distribution of x is given by

$$p(\boldsymbol{x} \mid \Omega^j_{t-1}, s_t) = \frac{p(s_t \mid \boldsymbol{x})p(\boldsymbol{x} \mid \Omega^j_{t-1})}{p(s_t \mid \Omega^j_{t-1})} \tag{3.1}$$

where the likelihood function is conditionally independent of the priors so that $p(s_t \mid \boldsymbol{x}) = p(s_t \mid \boldsymbol{x}, \Omega^j_{t-1})$ for all j. Both the prior and the posterior are N-dimensional probability simplices, i.e. they are N-dimensional vectors with elements summing to 1. The likelihood function for the realized signal $p(s_t \mid \boldsymbol{x}) \in (0,1)^N$ is also an N-dimensional object but it is not necessarily a simplex.[8] However, the elements in $p(s_t \mid \boldsymbol{x})$ are proportional to the ratio of the posterior and the prior probabilities of each corresponding outcome $x_n \in X$, i.e.

$$p(s_t \mid x_n) \propto \frac{p(x_n \mid \Omega^j_{t-1}, s_t)}{p(x_n \mid \Omega^j_{t-1})}. \tag{3.2}$$

Since $p(s_t \mid \Omega^j_{t-1})$ in (3.1) is simply a normalizing constant that ensures that the posterior probabilities over different states x_n sum to one, the ratio of the posterior and prior probabilities for each x_n are sufficient to characterize the realized signal. From here on, when we refer to a *signal*, we take that to mean the N-dimensional object proportional to $p(s_t \mid \boldsymbol{x})$. (Also, note that the label associated with the particular signal outcome is irrelevant.)

3.2. **Extracting the common signal by minimizing Kullback-Leibler divergence.** We saw in the previous paragraph that it is possible to back out an implied signal that in principle can account for the entire belief revision from $p(\boldsymbol{x} \mid \Omega^j_{t-1})$ to $p(\boldsymbol{x} \mid \Omega^j_t)$ for each forecaster j. However, we want to estimate the conditional distribution $p(s_t \mid \boldsymbol{x})$ of the common signal available to every forecaster. In general, such a signal will not be able to explain the entire belief revision of every forecaster in a given period, but we can estimate it by imposing that it should explain the maximum amount of the cross-section of belief

[7]While s_t denotes a generic signal in this subsection, with a slight abuse of notation, we will also use s_t to denote the common signal in what follows.

[8]That is, while each element of $p(s_t \mid \boldsymbol{x})$ is a probability, it is not generally the case that $\sum_{n=1}^{N} p(s_t \mid x_n) = 1$.

revisions. To make this notion operational, we need to be specific about what "maximum amount" means.

For a given signal s_t and prior distribution $p(\boldsymbol{x} \mid \Omega_{t-1}^j)$ we can compute the Kullback-Leibler divergence between the observed posterior distribution $p(\boldsymbol{x} \mid \Omega_t^j)$ and the hypothetical beliefs $p(\boldsymbol{x} \mid \Omega_{t-1}^j, s_t)$ a forecaster would have after updating to the signal s_t as

$$KL(\Omega_t^j; \Omega_{t-1}^j, s_t) = \sum_{n=1}^{N} p(x_n \mid \Omega_t^j) \log \left(\frac{p(x_n \mid \Omega_t^j)}{p(x_n \mid \Omega_{t-1}^j, s_t)} \right). \tag{3.3}$$

We can then define the estimated common signal $\widehat{s}_t$ as

$$p\left(\widehat{s}_t \mid \boldsymbol{x}\right) = \arg \min_{p(s_t|\boldsymbol{x}) \in (0,1)^N} \sum_{j=1}^{J} KL(\Omega_t^j; \Omega_{t-1}^j, s_t) \tag{3.4}$$

so that the estimate of the common signal s_t is the signal that minimizes the sum of KL-divergences between the cross-section of observed posteriors and the cross-section of the hypothetical beliefs forecasters would have if $\widehat{s}_t$ was the only piece of additional information available in period t.

3.3. **Extracting the residual individual signals by inverting Bayes rule.** We define the individual signal s_t^j of forecaster j as the signal, that when combined with the common signal and the forecaster j's observed prior, results in a posterior belief equal to his or her observed posterior in the SPF. It can be backed out by inverting Bayes rule as in (3.2) so that for each $x_n \in X$ we have

$$p(\widehat{s}_t^j \mid x_n) \propto \frac{p(x_n \mid \Omega_{t-1}^j, \widehat{s}_t, s_t^j)}{p(x_n \mid \Omega_{t-1}^j, \widehat{s}_t)}. \tag{3.5}$$

The procedure is illustrated for a hypothetical cross-section of two forecasters in Figure 3.1. The blue distributions in the far left and far right columns are, respectively, the observed prior and posterior beliefs that we obtain from the survey data. The top row corresponds to the beliefs and signals of the first forecaster, the bottom row to that of the second. The first step of the procedure is to find the common signal (gray, left-of-center column) such that the sum of the Kullback-Leibler divergences between the implied intermediate beliefs (green, center column) and the observed actual posteriors (blue, far right column) for each forecaster are minimized. The second step uses the inverted Bayes rule (3.2) to find the individual signals (gray, right-of-center column) so that when the updated hypothetical intermediate beliefs are updated, each forecasters' posterior coincides with the observed posteriors (blue, far right column).

3.4. **Realized signals vs signal structures.** If we can observe how forecasters beliefs about a given event evolve over τ consecutive periods, we can observe $\tau - 1$ updates of these beliefs and hence back out $\tau - 1$ common signals about the event, as well as $\tau - 1$ individual signals for each participating individual forecaster. The procedure described above allow us to identify the likelihood $p(s_t \mid \boldsymbol{x})$ of the realized signals up to a constant of proportionality $c_j = p(s_t \mid \Omega_{t-1}^j)^{-1}$ for each period where we can observe a belief revision about x. Since knowledge of any function proportional to the likelihood function, i.e. any function of the

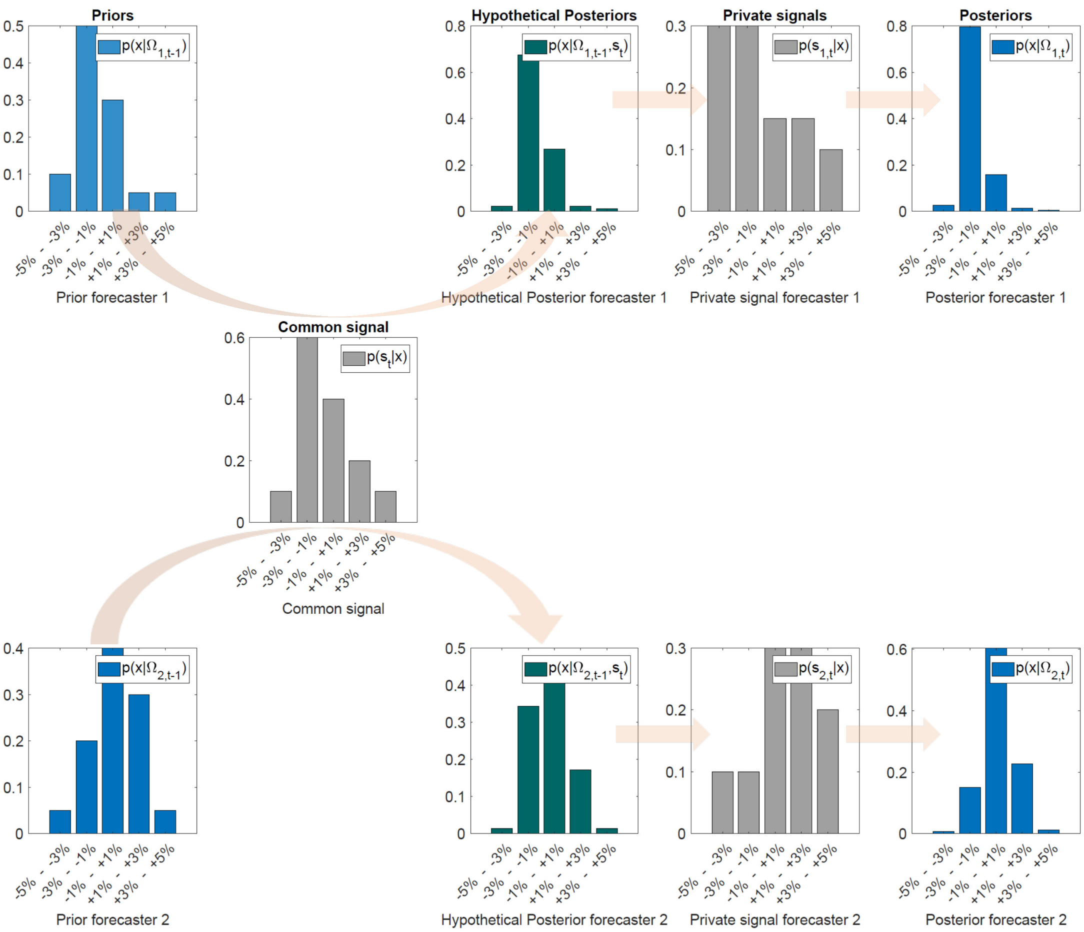

FIGURE 3.1. Illustration of procedure to estimate common and individual signals for a constructed example with $N = 5$. Beliefs are updated from left to right following the arrows. Blue graphs indicate observed beliefs. Gray graphs indicate signals. Green graphs indicate hypothetical intermediate beliefs implied by the priors and the common signal. The common signal is chosen to minimize the distance between the green intermediate beliefs and the observed posteriors in the right hand column. The individual signals are defined so that when the priors are updated with the common and respective individual signals, the implied posterior coincides with the observed posterior.

form $a \times p(s_t \mid \boldsymbol{x}) : a \in \mathbb{R}^+$, is sufficient to completely determine how agents update their prior in response to the signal s_t, extracting the properties of the realized signal up to a constant of proportionality is sufficient for our purposes. However, note that the procedure does not allow us to characterize the properties of forecasters' complete signal *structure*, i.e. it does not allow us to infer anything about the properties of other possible but unrealized signals $s_t \in S$, since doing so would require us to make assumptions about the invariance of the signal structure over time. Our procedure relies only on the information contained in the update between period $t-1$ and t to extract the signals in period t.

Note also that we do not impose that beliefs are rational in the sense of being model consistent or that forecast errors are not systematic. We simply impose that beliefs evolve *as if* agents apply Bayes' Rule to some, possibly mis-specified, probabilistic model.

Above we have used the language of information and signals to decompose the cross-section of beliefs revisions and this language naturally connects to the theoretical literature. In the next section, we use the first order condition for the minimization problem (3.4) to derive conditions on actual information structures that ensures that what the procedure extracts indeed corresponds to individual and common signals. There, we also characterize what the procedure finds under alternative modeling assumptions, for instance when different agents interpret a common signal differently or when the true underlying information structure is a standard linear Gaussian noisy rational expectations model.

4. Properties of the estimated signals and alternative information structures

The procedure above describes how the common and individual components of forecasters' information sets can be extracted from a panel of fixed-event probability forecasts. In this section we first provide a theoretical characterization of the signals that the proposed estimation procedure delivers in terms of sample moments. These results help our intuition for what the procedure will designate as common and individual information, and for how the mechanics of Bayesian updating determine how much and in what direction the extracted common component tilts beliefs.

The theoretical results presented here also allow us to characterize what the procedure finds under alternative information structures. In particular, we derive the theoretical counterparts to the estimated individual and common components in settings where (i) different agents interpret a common signal differently and (ii) the information structure is of a linear-Gaussian form commonly used in the theoretical imperfect information literature.

4.1. Properties of the estimated signals.

The estimated common signal is found by minimizing the sum of Kullback-Leibler divergences (3.4) between the beliefs it induces and the observed posteriors. In general, there is no common signal that will make these two sets of beliefs exactly equal for each forecaster. However, the following proposition characterizes the estimated common signal $\widehat{s}_t$ in terms of the posterior beliefs it induces relative to the observed average cross-section of posteriors.

Proposition 1. *The estimated common signal $\widehat{s}_t$ from 3.4 induces average beliefs equal to the average observed posterior distribution, i.e.*

$$\frac{1}{J}\sum_{j=1}^{J} p\left(x_n \mid \Omega_{t-1}^{j}, \widehat{s}_t\right) = \frac{1}{J}\sum_{j=1}^{J} p\left(x_n \mid \Omega_{t}^{j}\right) : n = 1, 2, ..., N. \tag{4.1}$$

Proof. In the Appendix. □

The proposition shows that the cross-sectional average beliefs induced by the common signal will in fact match the observed cross-sectional average posteriors. The logic of the proof is as follows. First, we show that the first order condition of (3.4) is sufficient to characterize the common signal $\widehat{s}_t$ that minimizes the sum of Kullback-Leibler divergences. The desired result then follows directly from manipulation of the first order condition. Corollary 1 characterizes the belief updates induced by the corresponding individual signals.

Corollary 1. *The estimated individual signals induce belief updates that average to zero across agents, i.e.*

$$\frac{1}{J}\sum_{j=1}^{J}\left[p\left(x_n \mid \widehat{s}_t^j, \widehat{s}_t, \Omega_{t-1}^j\right) - p\left(x_n \mid \widehat{s}_t, \Omega_{t-1}^j\right)\right] = 0 : n = 1, 2, ..., N. \tag{4.2}$$

The corollary follows simply from the fact that the average beliefs induced by the common signal matches the average observed posteriors. The average update to the individual signals must then average to zero across agents, since by construction, the individual signals are defined as the signals that, when combined with the common signal, induces the observed posterior beliefs.

Another way of understanding this result is to consider if, contrary to the corollary, the individual signals shifts the average posteriors towards a state x_n. This would imply that a different common signal s_t^* could achieve a smaller KL-divergence between the induced beliefs and the posteriors than the extracted common signal $\widehat{s}_t$ by setting $p\left(s^* \mid x_n\right) / \sum_{m=1}^{N} p\left(\widehat{s}_t \mid x_m\right)$ $> p\left(\widehat{s}_t \mid x_n\right) / \sum_{m=1}^{N} p\left(\widehat{s}_t \mid x_m\right)$. The signal $p\left(\widehat{s}_t \mid x\right)$ then cannot be the solution to the minimization problem (3.4).

The information in the common signal s_t influences the posterior only through the likelihood function $p(s_t \mid x)$. If the observed posteriors attach a higher average probability to state n than to state m relative to the observed priors, in order for Proposition 1 to hold, the extracted common signal must tilt beliefs towards state n relative to state m. To better understand what determines how much the common signal needs to favor one state over another, it is helpful to define the *mean-posterior-over-mean-prior odds ratio* of state n and m as follows.

Definition 1. The mean-posterior-over-mean-prior odds ratio R_m^n is defined as

$$R_m^n = \left(\frac{\frac{1}{J}\sum_{j=1}^{J} p\left(x_n \mid \Omega_t^j\right)}{\frac{1}{J}\sum_{j=1}^{J} p\left(x_m \mid \Omega_t^j\right)}\right) / \left(\frac{\frac{1}{J}\sum_{j=1}^{J} p\left(x_n \mid \Omega_{t-1}^j\right)}{\frac{1}{J}\sum_{j=1}^{J} p\left(x_m \mid \Omega_{t-1}^j\right)}\right) \tag{4.3}$$

The ratio R_m^n captures how much period t information shifts average beliefs in favor of state n relative to state m. As the following proposition shows, the ratio R_m^n serves as a baseline that captures how much more weight the extracted common signal puts on state n relative to state m in the special case when all agents share the same prior beliefs.

Proposition 2. *If the prior beliefs of all forecasters coincide, the relative probability of observing $\widehat{s}_t$ in states n and m equals the mean-posterior-over-mean-prior odds ratio so that*

$$\frac{p\left(\widehat{s}_t \mid x_n\right)}{p\left(\widehat{s}_t \mid x_m\right)} = R_m^n. \tag{4.4}$$

Proof. In the Appendix. □

In the special case of common priors, the relative likelihood of observing the common signal in states n and m are simply equal to how much more likely, on average, state n is perceived to be relative to state m after agents have observed period t information.

It is difficult (and we have been unable) to derive general results for when the equality (4.4) should be replaced with an inequality for arbitrary priors. However, the following two-agent example provides some intuition for why (4.4) fails to hold generally, and what determines the direction of the inequality that replaces it when agents have heterogeneous priors.

Proposition 3. *For $j \in \{1,2\}$, the extracted common signal puts more weight on state n relative to state m, compared to the common prior baseline so that*

$$\frac{p\left(\widehat{s}_t \mid x_n\right)}{p\left(\widehat{s}_t \mid x_m\right)} > R_m^n. \tag{4.5}$$

if the agent who a priori *thinks state n is relatively more likely, i.e.*

$$\frac{p\left(x_n \mid \Omega_{t-1}^1\right)}{p\left(x_m \mid \Omega_{t-1}^1\right)} > \frac{p\left(x_n \mid \Omega_{t-1}^2\right)}{p\left(x_m \mid \Omega_{t-1}^2\right)} \tag{4.6}$$

is also the agent who thinks the realized signal $\widehat{s}_t$ is more likely, i.e.

$$\frac{p\left(\widehat{s}_t \mid \Omega_{t-1}^1\right)}{p\left(\widehat{s}_t \mid \Omega_{t-1}^2\right)} > 1. \tag{4.7}$$

.

Proof. In the Appendix. □

The intuition for this result echoes the motivation of Shannon's (1948) definition of the quantity of information as the inverse of the probability of observing a signal. Bayesian updating implies that a less probable, and hence a more informative signal, changes beliefs more than a more probable signal, all else equal. A Bayesian agent will thus update his or her beliefs less, the higher his or her prior probability was of observing the realized signal. The proof follows from that this effect is concave in the probability of observing the signal. The signal then needs to be substantially more likely to be observed in state n than in state m, in order to move the prior beliefs of both agents sufficiently for the average posterior odds ratio of state n and m to shift by the factor R_m^n relative to the prior average odd ratio.

4.2. **Conditions for asymptotic convergence between extracted and true signals.** For discrete beliefs and information structures it is possible to derive conditions that ensures that the extracted common and individual signals asymptotically converge to the true signals as the number of agents become large.

Proposition 4. *Fix time t, let $p\left(s^j \mid x_n\right)$ be random variables with support $[0,1]$, and i.i.d. across j. The estimated signal converges in probability to the true common signal, i.e. $p\left(\widehat{s} \mid x_n\right) \xrightarrow{p} p\left(s \mid x_n\right)$ for all n as $J \to \infty$, if $p\left(s^j \mid x_n\right)$ is i.i.d. across n, and $p(x_n \mid s, \Omega_{t-1}^j) = \frac{1}{N}$ for every n and j.*

Proof. In the Appendix. □

To prove this result we treat the parameters of their likelihood functions associated with each forecasters individual signal as random variables. The conditions in the proposition ensure that the average update to the individual signal of the probability of each state averages to zero across agents. From Corollary 1, we know that this is a consequence of the first order condition (4.1), which in turn is sufficient to characterize the extracted signals.

The conditions in Proposition 4 are very stringent, and it is easy to think of settings where they are not satisfied. For instance, if all agents observe a perfectly precise individual signal, so that $p\left(s^j \mid x_m\right) = 0$ if $m \neq n$ for some $n \in \{1, 2, ..., N\}$, then the procedure will attribute the implied degenerate posteriors as being caused by a perfectly precise common signal. This example violates the condition that $E\left[p\left(s^j \mid x_n\right)\right] = E\left[p\left(s^j \mid x_m\right)\right]$.

A less extreme example is when the individual signals on average tilt beliefs towards some state x_n so that $E\left[p\left(s^j \mid x_n\right)\right] > E\left[p\left(s^j \mid x_m\right)\right]$ for $n \neq m$. There will then be a common component in the individual signals that will be attributed to the common signal by our procedure.

Finally, if the beliefs $p(x_n \mid s_t, \Omega_{t-1}^j)$ are random, or non-random but not uniform, the average belief update to the individual signals will not average to zero in the cross-section even if $E\left[p\left(s^j \mid x_n\right)\right] = E\left[p\left(s^j \mid x_m\right)\right]$ for all pairs of n and m. It is thus only under special circumstances that the individual and common signals can be interpreted as literally being different signals. In settings where these conditions are not satisfied, the appropriate interpretation is that the procedure extracts individual and common components of the new information available to agents in a given period that can be characterized *as if* they were single signals.[9]

4.3. **Different agents interpret a common signal differently.** In a rational expectations model, all agents have model consistent expectations and hence share the same model. In such a setting, all agents also interpret a common signal the same way. However, in a world where different agents may use different models, agents may use different likelihood functions to update their beliefs even to a common signal. To allow for different agents using different models, all we need to do is to treat the likelihood function each agent associates with the common signal as agent-specific, rather than the signal itself. Agent j's posterior is then given by

$$p_j(x \mid \Omega_{t-1}^j, s_t) = \frac{p_j(\widehat{s}_t \mid x)p(x \mid \Omega_{t-1}^j)}{p(s_t \mid \Omega_{t-1}^j)} \tag{4.8}$$

where the key notational difference is the j index on the likelihood function and the posterior. The special case with common priors is again helpful, since the extracted common signal is then a simple function of the cross-sectional averages of the agent specific likelihood functions.

Corollary 2. *With agent specific likelihood functions but a common prior, the estimated common signal satisfies*

$$\frac{p\left(\widehat{s}_t \mid x_n\right)}{p\left(\widehat{s}_t \mid x_m\right)} = \frac{\frac{1}{J}\sum_{j=1}^{J} p_j(s_t \mid x_n)}{\frac{1}{J}\sum_{j=1}^{J} p_j(s_t \mid x_m)} \tag{4.9}$$

for each pair $n, m \in 1, 2, ..., N$.

The proof follows from taking the ratio of the averages of (4.8) across agents for state n and m and then combining it with Proposition 1. In the case of heterogeneous priors, the

[9]Numerical simulations with a large number of forecasters indicate that signals that are independent across forecasters and satisfy $E\left[p\left(s^j \mid x_n\right)\right] = E\left[p\left(s^j \mid x_m\right)\right]$ generate estimated common signals that are numerically very close to the true signals if $p(x_n \mid s_t, \Omega_{t-1}^j)$ are also independent across forecasters.

expressions are again more complicated, but the logic and intuition of Propositions 2 and 3 above apply also to the case when a single common signal is interpreted differently.

4.4. **Linear Gaussian signal extraction.** Dating back to the classic papers of Lucas (1972), Grossman and Stiglitz (1976), Hellwig (1980) and Admati (1985), a large theoretical literature has used linear-Gaussian information structures to study economic decisions in settings where agents have private information.[10] The key advantage of this structure is its tractability, yielding closed form solutions for agents' posteriors. Given its continuing popularity and prominence in the theoretical imperfect information literature it is of interest to ask what our method would find if the observed survey data was generated by an underlying linear-Gaussian information structure. To study this question, we here first describe the standard linear-Gaussian set-up and what it implies for agents' beliefs.

Denote the prior beliefs of agent j as $x \mid \Omega_{t-1}^j \sim N\left(\underline{\mu}^j, \underline{\sigma}^2\right)$ and let the dispersion of prior means be normally distributed so that $\underline{\mu}^j \sim N\left(\underline{\mu}, \sigma_\mu^2\right)$. All agents observe the common signal s_t that is the sum of the true x and a common noise shock η

$$s_t = x + \eta : \eta \sim N\left(0, \sigma_\eta^2\right) \tag{4.10}$$

as well as an individual signal s_t^j of a similar form

$$s_t^j = x + \varepsilon^j : \varepsilon^j \sim N\left(0, \sigma_\varepsilon^2\right) \tag{4.11}$$

but where the noise shock ε^j is agent specific. The next lemma, which simply summarizes well-known results, characterizes the posterior beliefs implied by this structure.

In the linear-Gaussian information structure, the posterior of agent j is given by a Gaussian distribution such that

$$E\left(x \mid \Omega_{t-1}^j, s_t, s_t^j\right) = g_\mu \underline{\mu}^j + g_s s_t + g_j s_t^j \tag{4.12}$$

$$var\left(x \mid \Omega_{t-1}^j, s_t, s_t^j\right) = \left(\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}\right)^{-1} \tag{4.13}$$

where

$$g_\mu = \frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}}, g_s = \frac{\sigma_\eta^{-2}}{\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}}, g_j = \frac{\sigma_\varepsilon^{-2}}{\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}}. \tag{4.14}$$

The SPF data allows us to observe the prior $\underline{\mu}^j$ and posterior mean $E\left(x \mid \Omega_{t-1}^j, s_t, s_t^j\right)$ of each forecaster. How much the posterior mean shift relative to the prior depends on the realized values of both the common and individual signal. However, as the next proposition shows, observing a single cross-section of revisions does not by itself provide enough information to uniquely identify the realized values of both the common and individual signals.

Proposition 5. *In the linear-Gaussian information structure, there exist infinitely many tuples* $\left(s_t, s_t^1, \ldots, s_t^J\right)$ *that produce the same cross-section of posterior means* $\left(E\left(x \mid \Omega_{t-1}^1, s_t, s_t^1\right), \ldots, E\left(x \mid \Omega_{t-1}^J, s_t, s_t^J\right)\right)$.

Proof. In the Appendix. □

[10]See for instance the literature overviews in Veldkamp and Baley (2022) or Angeletos and Lian (2022).

The intuition of this result is simply that the cross-sectional average posterior mean can shift up (or down) relative to the average prior, either because the realized value of the common signal is higher (or lower) than expected, or because the average of the individual signals is higher (or lower) than expected.

If one were willing to assume a time-invariant stochastic structure of x, s_t and s_t^j, one could estimate the variances in the linear-Gaussian structure (4.10)-(4.14) from the panel of forecasts revisions. It would then be possible to assign probabilities to different realizations of signal tuples $\left(s_t, s_t^1, \ldots, s_t^J\right)$. Our proposed method does not require any assumptions about time-invariance of information structures. However, as the next proposition shows, our it still delivers a single unique estimate of the realized signal tuple with well-defined properties.

Proposition 6. *Up to the discrete approximation, the estimated common signal $\widehat{s}_t$ has conditional distribution*

$$\widehat{s}_t \mid x \sim N\left(x, \widehat{\sigma}_\eta^{-2}\right) \tag{4.15}$$

with estimated realized signal value given by

$$\widehat{s}_t = (1-\widehat{g})^{-1}\left[\left(g_\mu - \widehat{g}\right)\underline{\mu} + g_s s + g_j x\right] \tag{4.16}$$

where $\widehat{g} = \frac{\underline{\sigma}^{-2}}{\widehat{\sigma}_\eta^{-2}+\underline{\sigma}^{-2}}$ and $\widehat{\sigma}_\eta^{-2}$ solves the equation

$$g_\mu^2\sigma_\mu^2 + g_j^2\sigma_\varepsilon^2 + \left(\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}\right)^{-1} = \widehat{g}^2\sigma_\mu^2 + \left(\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2}\right)^{-1}. \tag{4.17}$$

Proof. In the Appendix. □

While algebraically somewhat involved, the proof strategy is conceptually quite simple. It only requires finding a realized $\widehat{s}_t$ and a likelihood function $p\left(\widehat{s}_t \mid x\right)$ such that the first order condition $\int_j p\left(x \mid \widehat{s}_t, \Omega_{t-1}^j\right) dj = \int_j p\left(x \mid s_t, s_t^j, \Omega_{t-1}^j\right) dj$ holds. That is, the first order condition implies that the estimated mean individual signal is such that the average cross-sectional mean is unchanged compared to if forecasters only observed the common signal. It is thus this condition that pins down a unique tuple of signals $\left(s_t, s_t^1, \ldots, s_t^J\right)$ among all tuples consistent with the cross-section of belief revisions.

It is possible to solve (4.17) for $\widehat{\sigma}_\eta^{-2}$ explicitly, but the resulting expression is not very informative. The following corollaries summarizes the key properties of $p\left(\widehat{s}_t \mid x\right)$ and $\widehat{\sigma}_\eta^{-2}$.

Corollary 3. *The estimated common signal $\widehat{s}_t$ coincides with s_t for all realizations if and only if $\sigma_\varepsilon^2 \to \infty$.*

The corollary states that it is only when the individual signals are completely uninformative that the extracted common signal generally coincides with the true realized signal. To understand this result, consider when both the common and individual signals are informative. Clearly, the common signal will shift the location of the average posterior. However, if the individual signals are informative, they will also shift the average posterior towards the true value of x. In order for the average posterior induced by the estimated common signal to coincide with the observed average, the estimated common signal need to shift beliefs in way that accounts for both the true common signal and the average shift towards the true value of x induced by the individual signals.

Corollary 4. *If the true common signal is uninformative* $\left(\sigma_\eta^2 \to \infty\right)$, *then the estimated common signal is of the form* $\widehat{s}_t = \alpha(x - \beta\underline{\mu})$ *with* $\alpha \geq 1$ *and* $\beta \leq 1$ *with estimated precision* $\widehat{\sigma}_\eta^{-2} < \sigma_\varepsilon^{-2}$.

If the true common signal is uninformative, the procedure will attribute the part of the shift in in the location of the posterior driven by the average individual signal as being caused by the estimated common signal.

Corollary 5. *The estimated precision* $\widehat{\sigma}_\eta^{-2}$ *is increasing in both* σ_ε^{-2} *and* σ_η^{-2} .

The corollary states that the precision of the estimated common signal is increasing in the precision of both the common and individual true signals. This implies that if the underlying information structure is Gaussian, our procedure will attribute increases in precision of the individual signals as partly being due to a more precise common signal. The formal proof is in the Appendix, but the result follows from the fact that both sides of the equation (4.17) are average posterior forecast errors. Both must be decreasing in the precision of all types of signals. From this, the results then follows from the implicit function theorem.

Corollary 6. *The estimated private signals* $\widehat{s}^j$ *have precision*

$$\widehat{\sigma}_\varepsilon^{-2} = \sigma_\varepsilon^{-2} - \left(\widehat{\sigma}_\eta^{-2} - \sigma_\eta^{-2}\right) \tag{4.18}$$

and sample mean given by

$$\int \widehat{s}^j dj = g_\mu \underline{\mu} + g_s s + g_j x. \tag{4.19}$$

The corollary follows from that when combined, the estimated individual and common signals imply a posterior that is equal to the true posterior. Since the individual signals cannot change the cross-sectional average distribution relative to the beliefs implied by the common signal, the mean individual signal must equal the average expected value of x conditional on the common signal. The variance in (4.18) is simply given by equating the posterior variance implied by s and s^j with the posterior variance implied by $\widehat{s}$ and $\widehat{s}^j$.

It is clear from the definition (3.4) that our procedure is designed to maximize the importance of the common signal. What we have shown here is what this implies for what the procedure finds under alternative underlying information structures. It is only under special conditions that what the procedure labels common and individual signals match with true underlying population objects with the same interpretation. The procedure thus provides an upper bound for the importance of the common signal, both in terms of how precise it is estimated to be and how much of the first-moment of the observed belief revisions that it can explain.

5. Three measures of signal informativeness

We want to quantify the informativeness of common and individual signals and study their cyclical properties. For this purpose, we here define three measures capturing different aspects of signal informativeness. To facilitate comparisons, each measure is defined so that a higher value indicates a more informative signal.

5.1. **The belief update measure.** A natural starting point is to find a measure that quantifies how much a given signal changes a prior belief. One such measure is the *belief update measure*, defined as the Kullback-Leibler divergence between the prior and posterior distributions.

Definition 5.1. *The belief update measure* $KL(\Omega_{t-1};\Omega_{t-1},s_t)$ *of the signal* s_t *is defined as*

$$KL(\Omega_{t-1};\Omega_{t-1},s_t)=\sum_{n=1}^{N}p(x_n\mid\Omega_{t-1}^{j})\log\left(\frac{p(x_n\mid\Omega_{t-1}^{j})}{p(x_n\mid\Omega_{t-1}^{j},s_t)}\right)\tag{5.1}$$

The belief update measure is large when the signal s_t results in a posterior distribution that is very different from the prior. From Bayes rule, this measure depends on how different the conditional signal probability ratios $p(s_t\mid x_n)/p(s_t\mid x_m)$ are from the corresponding prior ratios $p\left(x_n\mid\Omega_{t-1}^{j}\right)/p\left(x_m\mid\Omega_{t-1}^{j}\right)$. Hence, while it measures how much a signal affects the forecasters' beliefs, it depends not only on the signal but also on the forecaster's prior beliefs.

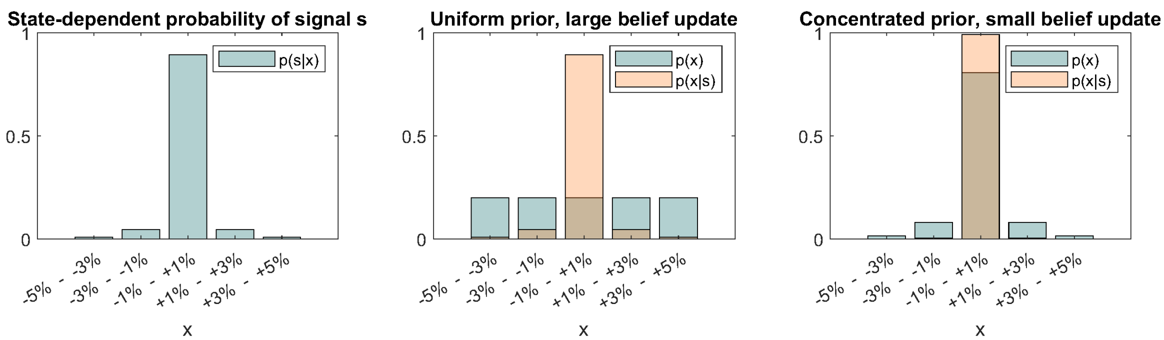

FIGURE 5.1. The signal s with concentrated conditional probabilities $p(s\mid x)$ (left panel) implies a large belief revision if the prior is uniform (middle panel) but a smaller revision if the prior is already concentrated (right panel), illustrating the dependence of the belief update measure on the prior distribution.

The role of the prior for the belief update measure is illustrated in Figure 5.1. The uniform prior (middle panel), when combined with a signal that has most of the mass in the tail regions of the distribution (left panel), implies a revision that re-allocates a lot of the mass from the tail regions towards the central bin. This results in a large Kullback-Leibler divergence between the prior and the posterior. A more concentrated prior (right panel) that already has most of the mass in the central bin, would be only marginally updated after the observation of the same signal leading to a correspondingly small Kullback-Leibler divergence between prior and posterior. A given signal will according to the update measure thus be considered more informative for a forecaster whose posterior changes a lot relative to his or her prior.

5.2. **The negative entropy measure.** Entropy is a measure of uncertainty, and for discrete distributions it is maximized by a uniform distribution. Entropy, and hence uncertainty, is minimized when there is only one possible outcome, i.e. for the degenerate distribution. We define the *negative entropy measure* as the negative of the posterior entropy a (hypothetical) agent with a uniform prior would have after having observed the signal.

Definition 5.2. *The negative entropy measure $H(s_t)$ of a signal s_t is defined as*

$$H(s_t) = \sum_{n=1}^{N} p\left(x_n \mid \Omega^u, s_t\right) \log p\left(x_n \mid \Omega^u, s_t\right) \tag{5.2}$$

where Ω^u denotes a uniform prior over outcomes in X.

The negative entropy measure is thus independent of forecasters' beliefs and a function only of the conditional signal probabilities $p(s \mid x)$. It captures the notion that a signal that is only likely to be observed in a specific state $x_n \in X$ is more informative than a signal that is equally likely to be observed in many states.

5.3. **The precision measure.** The belief update and the entropy measures are independent of the numerical labels associated with each outcome x_n and would remain unchanged if we reordered the outcome bins for the variable x. Hence, it does not distinguish between a signal that assigns all the probability mass to two central bins and a signal that assigns all the probability mass to two bins in the tails of the distribution. The *precision measure* remedies this and allows us to talk about the precision of a signal.

Definition 5.3. *The precision measure $P(s_t)$ of a signal s_t is defined as*

$$P(s) = var\left(x \mid \Omega^u, s_t\right)^{-1} \tag{5.3}$$

The measure $P(s_t)$ is thus the inverse of the variance of the posterior beliefs of a (hypothetical) agent with uniform prior have after having observed the signal. Defining it requires us to assign numerical values for each outcome $x_n \in X$. We do so by simply associating each interior outcome with the mid-point of the interval as defined in the SPF. For one-sided open boundary intervals we impose that the interval width is equal to the average interval length for the period. (Our results are robust to alternative ways to assign values to boundary intervals.)

Figure 5.2 illustrates how the entropy and the precision measure captures different aspects of the informativeness of a signal. Both signals imply that there are two bins that are much more likely than the remaining four bins, and both signals would be considered equally informative according to the entropy measure (5.2). However, the signal in the left panel assigns large weights to the two central bins, thus resulting in a high precision measure, while the signal in the right panel assign large weights to the boundary bins, resulting in a low precision measure.

One widely used measure of signal informativeness that is central to the large rational inattention literature that builds on the formalism proposed in Sims (1998, 2003), is *mutual information*. The mutual information $I(S;X)$ between two random variables S and X measures how much the entropy of X is reduced by observing S. Computing it requires knowledge of the entire conditional distribution $p(S \mid X)$, i.e. we would need to know $p(s_m \mid x_n)$ for each $s_m \in S$ where $m \in \{1, 2, ..., M\}$ indexes the labels of different signal realizations. As discussed above, estimating the entire signal structure would require imposing additional

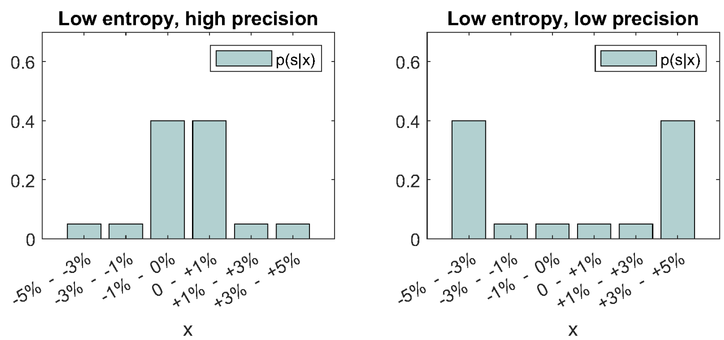

Figure 5.2. Illustration of the difference between entropy and variance measures of signal informativeness. From a uniform prior, the conditional probabilities of the signals in the left and right panel imply posterior distributions with the same entropy since the entropy measure is independent of the ordering of the bins. However, the signal in the left hand panel is more informative according to the precision measure.

restrictions on its time invariance and we do not pursue this in the current paper. Below we will use the measures defined here to quantify how signal informativeness differs over time, across variables and across individual and common signals.

6. Empirical properties of individual and common Information

In this section we apply the procedure described above to SPF probability forecasts. As an informal validation exercise for the methodology, we start by reporting how the measures of informativeness change in response to known macroeconomic events. We then study the cyclical properties of both common and individual information and how the two types of information covary with macroeconomic variables, forecasters' subjective probability of a recession, NBER dated recessions and with an index of expected stock market volatility. Finally, we document the relative informativeness of common and individual signals as well as the cross-sectional heterogeneity of signal informativeness across forecasters.

Throughout this section, our focus is on the current year forecasts. As shown by Bassetti, Casarin and Del Negro (2022), at short horizons, the SPF survey respondents' perceived precision of their forecasts corresponds closely to their actual forecast accuracy, while at longer horizons, the relationship is more tenuous or non-existent. Our measures of the subjective precision of signals then also translates into actual precision in forecasting at short horizons.[11] Whether the distinction that we measure forecasters' subjective uncertainty matters or not depends on the purpose of the analysis. For instance, if we are interested in whether forecasters form beliefs rationally as in Bassetti, Casarin and Del Negro (2022), then it certainly matters whether forecasts are more accurate when they perceive them to be so.

[11]Results are qualitatively similar for longer forecasts horizons, though there is generally less variation in signal informativeness about next and the-year-after-next calendar year outcomes. The complete results are available through the replication files.

However, for understanding decisions taken under uncertainty, it is arguably the subjective uncertainty that is more relevant.

In order to interpret the results below appropriately, it is worth remembering that we proceed *as if* the common component in belief revisions is caused by a single common signal and *as if* any residual individual revision is caused by a forecaster specific signal. We extract the common and individual signals for each forecaster at each quarter about current year outcomes of CPI inflation, unemployment, GDP growth, GDP deflator and PCE inflation. The procedure thus associates an N-dimensional vector of probabilities with each type of signal at each point in time. As a summary of the informativeness of the common and individual signals and how it varies over time, we compute the time series of each measure of informativeness for the common signal together with the cross-sectional average informativeness of the individual signals. The time-series of these measures for CPI inflation, unemployment and GDP growth are plotted in Figure 6.1.[12]

6.1. **Signal informativeness and known macroeconomic events.** Information about the macro economy does not exist in a vacuum, but is generated and acquired jointly with economic outcomes. The incentives for economic agents to acquire information about the economy is likely to depend on economic conditions, e.g. Flynn and Sastry (2022). Economic conditions also affect how much news media focus on the economy, e.g. Nimark (2014) and Chahrour, Nimark and Pitschner (2021) as well as how much and in what manner central banks communicate about the economy, e.g. Herbert (2021). We would thus a priori expect that our measures of informativeness should increase in response to major macroeconomic events. As an informal validation check, we first document that this is indeed the case.

Figure 6.1 shows that signal informativeness varies substantially over time and that the variation tends to be clustered in time. The sample period includes two major macroeconomic events, the financial crisis/Great Recession of 2008-2009 and the onset of the COVID pandemic in the second quarter of 2020. A common pattern across variables and measures is that signal informativeness increases during both the Great Recession and during COVID. That this pattern is present also in the precision measure suggests that, while these were periods of high macroeconomic volatility, they were not necessarily periods of high perceived uncertainty.

Some changes in informativeness are specific to some macro variables and events. For CPI inflation, the financial crisis initially resulted in a larger increase in the informativeness of the common signal. As demonstrated in the previous section, an increase in the informativeness of the common signal could be driven by either an actual increase in its informativeness or by a more dominant common component in the cross-section of individual signals. In the period following the acute phase of the financial crisis, individual signals explain more of forecasters' belief revisions and they are also perceived to have increased in precision. However, since this perceived increase in precision is attributed to the individual signal, there must also be an increase in the cross-sectional differences in beliefs, since by Corollary 6 the update to the individual signal must average to zero in the cross-section. Furthermore, since the same episode is associated with an increased precision also of the common signal, one should be

[12]The corresponding graphs for GDP deflator, PCE inflation and the entire sample for GDP growth are reported in the Appendix.

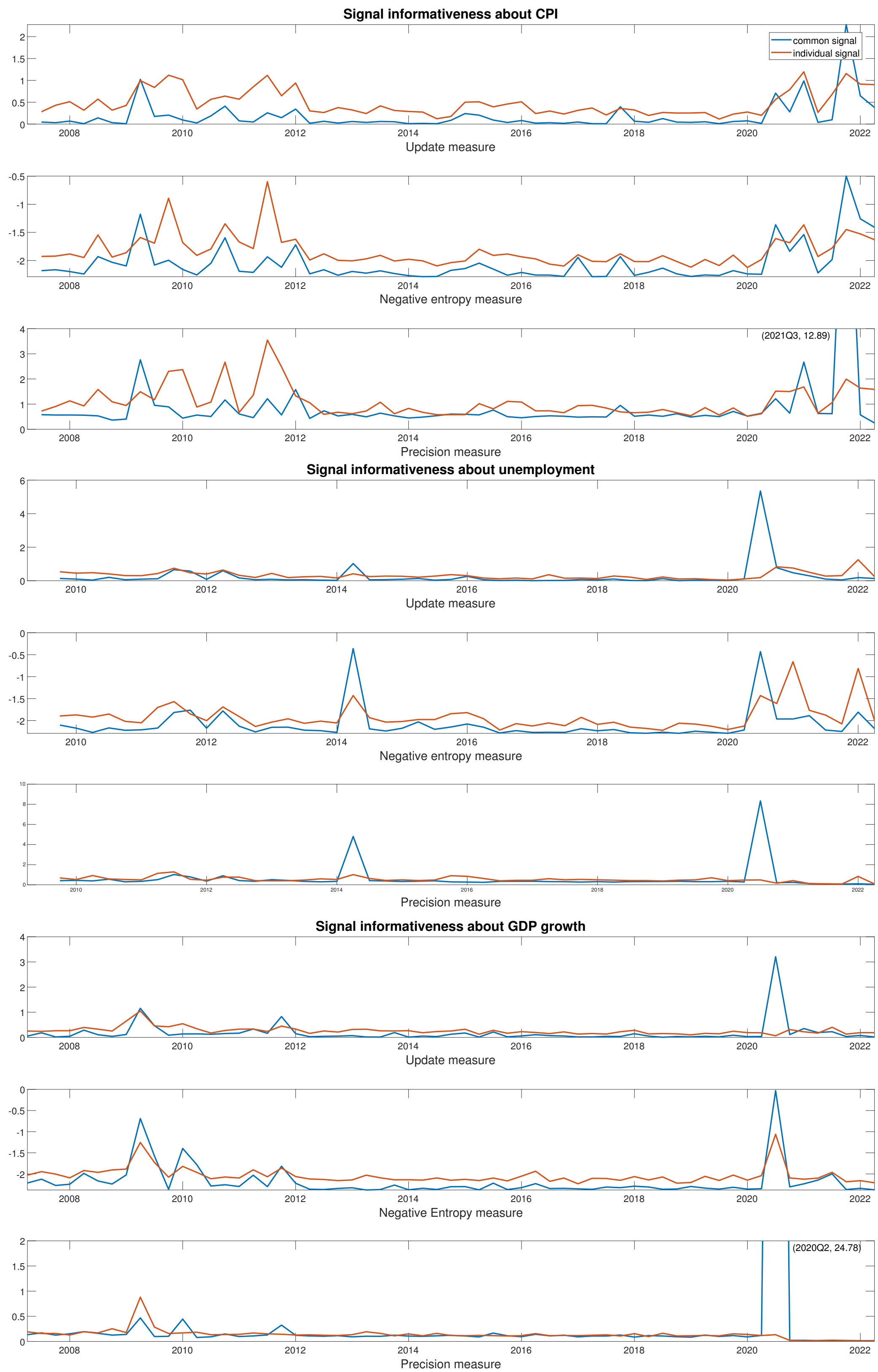

FIGURE 6.1. Time series of informativeness of individual and common signals about CPI inflation, unemployment and GDP growth.

cautious to interpret the increase in cross-sectional dispersion of beliefs as an increase in forecasters' subjective uncertainty, e.g. Bachmann, Elstner, and Sims (2013).

The initial stages of the COVID pandemic is associated with an increase in the informativeness of both the individual and the common signal about CPI inflation. However, there is a very sharp increase in the informativeness of the common signal coinciding with the sharp increase in actual CPI inflation mid-way through 2021. That this period is associated with a sharp increase in the informativeness of the common signal generally is likely a consequence of the intense media focus on inflation at the time.

For the informativeness of the signals about unemployment, the time around the financial crisis does not stand out in the same way as it does for CPI inflation. However, the informativeness of the common signals about unemployment did increase sharply during the COVID pandemic. Unlike for inflation though, this increase happened early in the pandemic, perhaps reflecting that the increase in actual unemployment occurred much faster and more dramatically than the increase in CPI inflation that occurred only towards the tail end of the pandemic.

One interesting episode that the procedure picks up occurs in 2014:Q2. If one were to only look at the unemployment rate outcome in that quarter, there is nothing to suggest that anything special was going on. However, the Federal Reserve had previously stated that they would leave interest rates at or near 0% until unemployment fell below 6.5%, see Federal Reserve Board of Governors (2012). This threshold was crossed in 2014:Q2 and because of its significance for future policy changes, it received a large amount of media attention, e.g. New York Times (2012). For GDP growth, we can see that as for the other variables, the informativeness of signals increases during both the financial crisis and during COVID. The procedure thus captures the kind of events that we a priori would expect to influence how much information agents acquire about macroeconomic variables.

6.2. **The cyclical properties of signal informativeness.** To document the cyclical properties of signal informativeness, we here first compute the correlations between our measures and the associated underlying macro variables. We also compute the correlations of our measures with the subjective probability of a recession, actual NBER dated recessions and the VIX measure of expected stock market volatility implied by option prices.

Correlations between signal informativeness and macro variables. Table 1 reports the correlation between the informativeness of signals and outcomes for CPI inflation as well as the correlation with lagged outcomes and measures of volatility, i.e. the magnitude of absolute changes. High inflation tends to be associated with more informative common signals, though this pattern is generally not statistically significant except for the negative entropy measure and only at the 10% level. High inflation is negatively uniformly negatively correlated with the informativeness of individual signals, but statistically significantly so only for lagged inflation and then only at the 10% level for the negative entropy and precision measures.

The strongest, and statistically most significant correlations overall, are between the informativeness of signals and lagged absolute changes in inflation, suggesting that forecasters become more informed from both individual and common sources when inflation is more volatile. However, as shown in the bottom panel of Table 1, when the COVID period

2020:Q2-2023:Q2 is excluded from the sample, practically all the correlations between the informativeness of the common signal and the inflation outcomes become statistically insignificant.

Excluding the COVID period also makes the correlations between the informativeness of the individual signals and the level of inflation uniformly significant at the 1% level. These correlations are strongly negative at around -0.45 This suggests that in the non-COVID part of the sample, there is a strong cyclical component in forecasters' acquisition of individual information associated with the level of inflation.It is also interesting that it is only when the COVID period is included in the sample that the common component in forecasters' information sets becomes more important. This is consistent with the model and evidence presented in Nimark and Pitschner (2019), where major events, such as the COVID inflation surge, leads to more homogeneous news coverage.

CPI inflation

	π_t^{cpi}	π_{t-1}^{cpi}	$\Delta\pi_t^{cpi}$	$\lvert\Delta\pi_t^{cpi}\rvert$	$\lvert\Delta\pi_{t-1}^{cpi}\rvert$
Individual signals					
KL	-0.08	-0.13	0.08	0.48***	0.45***
H	-0.20	-0.22*	-0.03	0.36***	0.35***
P	-0.17	-0.22*	0.05	0.36***	0.35***
Common signals					
KL	0.12	0.15	-0.03	0.23*	0.44***
H	0.25*	0.21*	0.14	0.45***	0.53***
P	0.02	0.04	-0.12	-0.06	0.29**

CPI inflation (excluding COVID)

	π_t^{cpi}	π_{t-1}^{cpi}	$\Delta\pi_t^{cpi}$	$\lvert\Delta\pi_t^{cpi}\rvert$	$\lvert\Delta\pi_{t-1}^{cpi}\rvert$
Individual signals					
KL	-0.45***	-0.13***	-0.16	0.47***	0.52***
H	-0.44***	-0.38***	-0.13	0.38***	0.47***
P	-0.44***	-0.41***	-0.05	0.42***	0.46***
Common signals					
KL	-0.17	-0.12	-0.14	0.16	0.13
H	-0.14	-0.09	-0.13	0.24*	0.19
P	-0.10	-0.06	-0.12	0.17	0.19

Table 1. Correlation of CPI inflation outcomes and associated information measures, with and without including the COVID sample. Asterisks *,** ,*** denote 10%, 5%, and 1% significance levels.

The top panel of Table 2 reports that high (lagged) levels of unemployment tend to be associated with both individual and common signals being more informative and precise. Interestingly, while high levels of unemployment tend to be associated with more informative signals, increases in unemployment tend to be associated with less informative and less precise

Unemployment

	u_t	u_{t-1}	Δu_t	$\lvert\Delta u_t\rvert$	$\lvert\Delta u_{t-1}\rvert$
			Individual signals		
KL	0.27*	0.38***	-0.18	-0.06	-0.19
H	0.16	0.31**	-0.24*	0.07	-0.10
P	0.32**	0.28**	0.06	-0.11	-0.11
			Common signals		
KL	0.22	0.48***	-0.41***	0.38***	0.14
H	0.20	0.40***	-0.31**	0.24*	0.04
P	0.21	0.43***	-0.35***	0.31**	0.12

Unemployment (excluding COVID)

	u_t	u_{t-1}	Δu_t	$\lvert\Delta u_t\rvert$	$\lvert\Delta u_{t-1}\rvert$
			Individual signals		
KL	0.73***	0.73***	-0.08	0.13	0.14
H	0.50***	0.51***	-0.18	0.16	0.09
P	0.36**	0.35**	0.05	-0.07	0.04
			Common signals		
KL	0.31**	0.32**	-0.24	0.24	0.15
H	0.18	0.20	-0.31**	0.35**	0.08
P	0.11	0.14	-0.32**	0.37**	0.02

Table 2. Correlation of unemployment outcomes and associated information measures, with and without including the COVID sample. Asterisks *,**,*** denote 10%, 5%, and 1% significance levels..

signals. This pattern also holds, but becomes somewhat less statistically significant when the COVID sample is excluded (bottom panel). One possible interpretation of this result is that in a recession, when unemployment increases rapidly, there is an increased uncertainty about how high unemployment will go before it peaks.

As with the signals about CPI inflation, excluding the COVID sample increases both the magnitude and the statistical significance of the correlations between the informativeness of the individual signals and the level of unemployment. These correlations now range from 0.36 to 0.73 and they are all significant at the 5% or 1% level.

Signal informativeness and recessions. The Federal Reserve Bank of Philadelphia, who administers the SPF, computes the so-called *Anxious Index* which measures the SPF respondents' subjective probability of a recession. The survey asks panelists to estimate the probability that real GDP will decline in the quarter in which the survey is taken and in each of the following four quarters. The anxious index is the average reported probability of a decline in real GDP in the quarter after a survey is taken. As shown in Table 3, for almost all measures and variables, this index is positively correlated with each of our measures of signal informativeness, and for a majority of the measures, these correlations are statistically significant.

	CPI inflation	*unemployment*	*GDP growth*	*GDP deflator*	*PCE inflation*
			Individual signals		
KL	0.20	0.06	0.27***	0.23***	0.24*
H	0.15	0.24*	0.27***	0.17**	0.24*
P	0.13	-0.20	-0.02	-0.06	0.23*
			Common signals		
KL	0.16	0.72***	0.18***	0.08	0.19
H	0.26*	0.45***	0.24***	0.14*	0.17
P	0.03	0.58***	0.15**	-0.10	0.04

Table 3. Correlation between the Philadelphia Fed's *Anxious Index* and the measures of informativeness. *,** ,*** denote 10%, 5%, and 1% significance levels respectively.

The *Anxious Index* captures forecasters' subjective probabilities of a recession. It is also the subjective probability of a recession that should matter for incentives for acquiring more precise information. The correlations between the informativeness of the signals and actual NBER dated recessions are generally weaker, and the sign of these correlations is not uniform across measures and variables. There does thus not appear to be any systematic relationship between the informativeness of signals and actual recessions.[13]

Signal informativeness and stock market volatility. The *VIX Index* is a popular measure of market expectations of the volatility of stock prices derived from prices of S&P 500 index options. It is provided by the Chicago Board Options Exchange. As shown in Table 4, with the exception of the precision of signals about inflation in the GDP deflator index, all measures of signal informativeness are positively correlated with the *VIX Index.* This indicates that when there is a high level of perceived uncertainty about the stock market, the incentives to acquire information by the survey participants are also particularly strong. Again, for a majority of the measures, these correlations are statistically significant.

There exists theoretical models that predict that recessions are times of reduced signal informativeness if economic activity by itself help generate information, e.g. Chalkley and Lee (1998), Veldkamp (2005), Van Nieuwerburgh and Veldkamp (2006), Ordoñez (2013), Fajgelbaum, Shaal and Taschereau-Dumouchel (2017). Other models, e.g. Chiang (2022), Song and Stern (2022) and Flynn and Sastry (2022), instead point point out that incentives to acquire information is stronger during recessions when the marginal utility of consumption is high and mistakes are more costly. Flynn and Sastry (2022) also find that empirically, firms that pay more attention to macroeconomic variables in recessions make smaller mistakes in hiring. This is consistent with our evidence that signal informativeness is positively correlated with the Philadelphia Fed's *Anxious Index* as well as with the *VIX Index* from the Chicago Board Options Exchange. The fact that the informativeness of the signals co-move

[13]The correlations with NBER dated recessions is available through the replication files.

	CPI inflation	*unemployment*	*GDP growth*	*GDP deflator*	*PCE inflation*
			Individual signals		
KL	0.29^{**}	0.36^{***}	0.25^{***}	0.12	0.22^{*}
H	0.29^{**}	0.30^{**}	0.20^{**}	0.10	0.23^{*}
P	0.32^{**}	0.03	0.17^{*}	-0.02	0.19
			Common signals		
KL	0.12	0.26^{*}	0.22^{**}	0.15^{*}	0.17
H	0.25^{**}	0.16	0.22^{**}	0.12	0.22^{*}
P	0.02	0.10	0.08	-0.07	0.05

Table 4. Correlation between *the VIX Index* and measures of informativeness. *,**,*** denote 10%, 5%, and 1% significance levels respectively.

positively with the *VIX Index* is also not surprising considering that the SPF participants are to large extent drawn from banks and other financial firms. It makes sense that the value of more precise information for such firms should be increasing in the perceived uncertainty of financial markets.

6.3. **Cross-sectional heterogeneity in signal informativeness.** To evaluate whether individual or common signals are on average more informative, we first compute the time-average of the informativeness of the common signal. We then compare this to the time-average informativeness of the forecasters' individual signals. The result of this is illustrated in Figure 6.2, where have plotted the average informativeness of the common signal (vertical red line) together with the histogram of the cross-section of the time-average informativeness of forecasters' individual signals. The labels on each panel indicate for each measure and each variable the fraction of forecasters who observe individual signals that are more informative than the common signal.

From the figure it is clear that, for all variables and measures except the precision measure for unemployment and GDP, a majority of forecasters observe individual signals that are more informative than the common signal. Without the spikes in precision associated with unemployment crossing the 6.5% threshold in 2014:Q2 and at the onset of COVID in 2020:Q2, this would hold also for the precision measure for unemployment. Similarly, this would also hold for the precision measure of GDP growth without the spike in 2020:Q2. The fact that for most measures and most forecasters, the individual signals are more informative is particularly noteworthy since the procedure to estimate the common signal by construction maximizes its importance. Our estimates can thus be interpreted as a lower bound of the informativeness of individual signals..

We are unaware of any papers that have focused on empirically estimating the relative informativeness of common and individual signals. However, there exists a small number of structural models that feature both private and public signals, and that have either been calibrated or estimated to match the dynamics of macroeconomic aggregates, e.g. Lorenzoni (2009) and Nimark (2014). While not the main focus of these papers, the relative informativeness of common and individual (or public and private) signals are addressed indirectly.

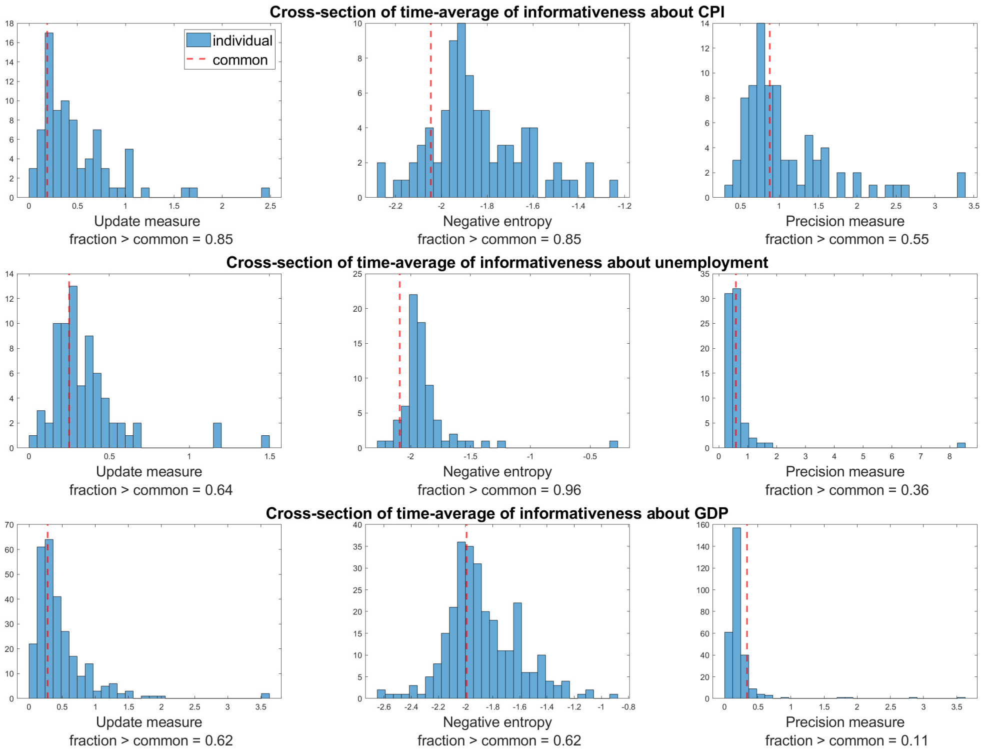

FIGURE 6.2. Cross-section of informativeness of individual signals relative to the common signal. The labels on each panel indicate for each measure and each variable the fraction of forecasters who observe individual signals that are more informative than the common signal.

A careful quantitative comparison of our results with the calibrated or estimated parameters from the structural macro literature would be quite involved, but these papers have generally used (or found) parameter values that imply that agents' private signals are more informative than the public signals. Our results are thus qualitatively consistent with the relative precision of the public and private signals in these papers.[14]

While the individual signals appear to both be more precise and explain more of forecasters belief revisions than the common signals, there is substantial cross-sectional heterogeneity. To get a sense of the magnitude of this heterogeneity, we can translate the precision measure into a measure of uncertainty by computing the implied posterior standard deviation of a hypothetical agent with a uniform prior who have observed a typical individual signal.

[14]One reason such a comparison is not straightforward is that the type of models used in Lorenzoni (2009) and Nimark (2014) impose substantial structure on the data that implies that there may be a feedback from the dynamics of macro aggregates to the estimated parameter values. Another issue that would complicate such a direct comparison is that agents in these models observe multiple private signals, that these signals are partly about exogenous variables, and that a single signal is to different degrees informative about several different endogenous variables.

The cross-sectional range between the 5th and 95th percentile of this measure is 0.64-1.42% for CPI inflation, 0.92-1.71% for unemployment and 1.47-3.83% for GDP growth. There is little existing theoretical work that either attempts to explain heterogeneity in information precision, or to study its consequences, with one interesting exception being Broer, Kohlhas, Mitman and Schlafman (2022).[15]

6.4. **The correlation between individual and common signal informativeness.** The correlation between the cross-sectional average informativeness of the individual signals and the informativeness of the common signal are positive for each variable and each measure. However, for the reasons explained in Section 4 above, this could be an artifact of the estimation procedure. However, the positive correlation between the informativeness of the common signal and the average informativeness of the individual signals masks substantial heterogeneity. As illustrated in Figure 6.3, for each measure and for each variable, a substantial fraction of forecasters observe individual signals whose informativeness is negatively correlated with that of the common signal. The labels on the panels indicate the fraction of forecasters who observe individual signals with an informativeness that co-varies negatively with the informativeness of the common signal. Again, since the method may produce a positive bias of these correlations, this is evidence that for at least some forecasters, more informative common information sources may lead them to substitute away from individual information sources.

Most of the theoretical literature that has studied endogenous information acquisition while allowing for both private and public information have found that more accurate public information crowds out individual information acquisition, e.g. Wong (2008) and Colombo, Femminis and Pavan (2014). Such a mechanism would suggest a negative correlation between the informativeness of individual and common signals. This mechanism thus appears to be able to account for the behavior of at least some of the forecasters.

7. Conclusions

In this paper we have proposed a method to decompose a panel of belief revisions *as if* driven by individual and common signals while imposing only weak assumptions. We showed theoretically how the properties of the estimated common and individual signals map into sample moments, which facilitates interpreting the implications of our empirical findings for alternative specific and more parametric underlying information structures.

When applied to probability forecasts from the *Survey of Professional Forecasters,* we find that (i) the informativeness of both common and individual signals is state-dependent, with major macroeconomic events, volatile inflation, high unemployment, perceived stock market volatility and a high risk of recession all tending to be associated with more informative signals, (ii) when excluding the COVID pandemic period, the cyclicality of signal informativeness is most pronounced in individual signals while the COVID pandemic is associated with a large increase in the informativeness of common signals across all variables.

[15]Some of the heterogeneity in individual signal informativeness is driven by forecasters entering and exiting the sample at different points in time. When we control for within-period cross-sectional mean informativeness, the standard deviation across forecasters fall by between 3 and and 30 per cent, depending on variable and measure.

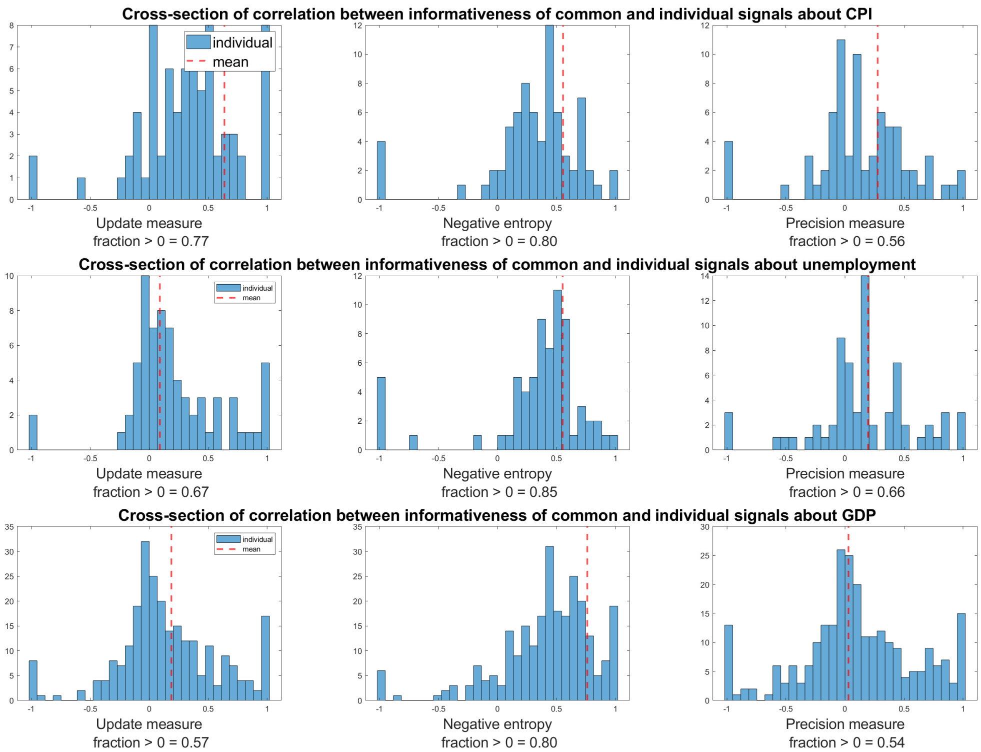

FIGURE 6.3. Cross-section of correlations between informativeness of individual and common signals. The labels on the panels indicate the fraction of forecasters who observe individual signals with an informativeness that co-varies negatively with the informativeness of the common signal.

The documented time variation and cross-sectional heterogeneity of signal informativeness presents a challenge to models that presume time-invariant and agent-homogeneous information structures. While the time-variation in the the informativeness of the common signals are qualitatively consistent with the model and evidence in Nimark and Pitschner (2019), more work is clearly needed to understand the macroeconomic consequences of time variation in the relative importance of the common component in forecasters belief revisions and the cross-sectional heterogeneity of individual signal informativeness. Future research may for use our decomposition to discipline theoretical models of how beliefs respond to major news events.

REFERENCES

[1] Admati, A.R., 1985. A noisy rational expectations equilibrium for multi-asset securities markets. Econometrica.

[2] Amador, M. and Weill, P.O., 2010. Learning from prices: Public communication and welfare. Journal of Political Economy, 118(5), pp.866-907.

[3] Amador, M. and Weill, P.O., 2012. Learning from individual and public observations of others' actions. Journal of Economic Theory, 147(3), pp.910-940.
[4] Angeletos, G.M., Iovino, L. and La'O, J., 2016. Real rigidity, nominal rigidity, and the social value of information. American Economic Review, 106(01), pp.200-227.
[5] Angeletos, G.M. and La'O, J., 2010. Noisy business cycles. NBER Macroeconomics Annual, 24(1), pp.319-378.
[6] Angeletos, G.M. and La'O, J., 2013. Sentiments, Econometrica, 81(2), pp.739-779.
[7] Angeletos, G.M. and Lian, C., 2023. Dampening general equilibrium: incomplete information and bounded rationality. In Handbook of Economic Expectations (pp. 613-645). Academic Press.
[8] Bachmann, R., Elstner, S. and Sims, E.R., 2013. Uncertainty and economic activity: Evidence from business survey data. American Economic Journal: Macroeconomics, 5(2), pp.217-249.
[9] Baley, I. and Veldkamp, L., 2023. Bayesian learning. In Handbook of Economic Expectations (pp. 717-748). Academic Press.
[10] Barillas, F. and Nimark, K.P., 2017. Speculation and the term structure of interest rates. The Review of Financial Studies, 30(11), pp.4003-4037.
[11] Barillas, F. and Nimark, K., 2019. Speculation and the bond market: An empirical no-arbitrage framework. Management Science, 65(9), pp.4179-4203.
[12] Bassetti, F., Casarin, R. and Del Negro, M., 2022. A Bayesian Approach to Inference on Probabilistic Surveys (No. 1025). Federal Reserve Bank of New York.
[13] Bonham, C.S. and Cohen, R.H., 2001. To aggregate, pool, or neither: Testing the rational-expectations hypothesis using survey data. Journal of Business & Economic Statistics, 19(3), pp.278-291.
[14] Bonham, C.S. and Dacy, D.C., 1991. In search of a " strictly rational" forecast. The Review of Economics and Statistics, pp.245-253.
[15] Born, B., Enders, Z., Menkhoff, M., Müller, G.J. and Niemann, K., 2023. Firm expectations and news: Micro v macro (No. 43). Working Papers of the Priority Programme 1859" Experience and Expectation. Historical Foundations of Economic Behaviour".
[16] Broer, T., A. Kohlhas, K. Mitman and K. Schlafman, 2022, Expectations and Wealth Heterogeneity in the Macroeconomy, working paper.
[17] Chahrour, R., Nimark, K. and Pitschner, S., 2021. Sectoral media focus and aggregate fluctuations. American Economic Review, 111(12), pp.3872-3922.
[18] Chalkley, M. and Lee, I.H., 1998. Learning and asymmetric business cycles. Review of Economic Dynamics, 1(3), pp.623-645.
[19] Chiang, Y.T., 2022. Attention and Fluctuations in Macroeconomic Uncertainty. FRB St. Louis Working Paper, (2022-4).
[20] Clements, M.P., 2006. Evaluating the Survey of Professional Forecasters probability distributions of expected inflation based on derived event probability forecasts. Empirical Economics, 31(1), pp.49-64.
[21] Clements, M.P., 2018. Are macroeconomic density forecasts informative?, International Journal of Forecasting, 34(2), pp.181-198.
[22] Clements, M.P., Rich, R.W. and Tracy, J., 2025. An investigation into the uncertainty revision process of professional forecasters. Journal of Economic Dynamics and Control, 173, p.105060.
[23] Coibion, O. and Gorodnichenko, Y., 2015. Information rigidity and the expectations formation process: A simple framework and new facts. American Economic Review, 105(8), pp.2644-78.
[24] Coibion, O., Gorodnichenko, Y. and Kumar, S., 2018. How do firms form their expectations? new survey evidence. American Economic Review, 108(9), pp.2671-2713.
[25] Coibion, O. and Gorodnichenko, Y., 2012. What can survey forecasts tell us about information rigidities?. Journal of Political Economy, 120(1), pp.116-159.
[26] Coibion, O., Gorodnichenko, Y. and Ropele, T., 2020. Inflation expectations and firm decisions: New causal evidence. The Quarterly Journal of Economics, 135(1), pp.165-219.
[27] Colombo, L., Femminis, G. and Pavan, A., 2014. Information acquisition and welfare. The Review of Economic Studies, 81(4), pp.1438-1483.

[28] Croushore, D.D., 1993. Introducing: the survey of professional forecasters. Business Review-Federal Reserve Bank of Philadelphia, 6, p.3.
[29] Croushore, D. and Stark, T., 2019. Fifty years of the survey of professional forecasters. Economic Insights, 4(4), pp.1-11.
[30] Diebold, F.X., Tay, A. and Wallis, K., 1997. Evaluating density forecasts of inflation: the Survey of Professional Forecasters, NBER Working Paper 6228.
[31] Engelberg, J., Manski, C.F. and Williams, J., 2009. Comparing the point predictions and subjective probability distributions of professional forecasters. Journal of Business & Economic Statistics, 27(1), pp.30-41.
[32] Fajgelbaum, P.D., Schaal, E. and Taschereau-Dumouchel, M., 2017. Uncertainty traps. The Quarterly Journal of Economics, 132(4), pp.1641-1692.
[33] Federal Reserve Board of Governors, 2012, Federal Reserve press release, https://www.federalreserve.gov/newsevents/pressreleases/monetary20121212a.htm .
[34] Financial Times, 2021, Fed's Powell warns inflationary supply chain snags may persist, https://www.ft.com/content/90fc98ad-d69b-44c5-8902-b93c4f952805.
[35] Flynn, J.P. and Sastry, K., 2024. Attention cycles (No. w32553). National Bureau of Economic Research.
[36] Ganics, G., Rossi, B. and Sekhposyan, T., 2024. From Fixed-Event to Fixed-Horizon Density Forecasts: Obtaining Measures of Multihorizon Uncertainty from Survey Density Forecasts. Journal of Money, Credit and Banking, 56(7), pp.1675-1704.
[37] Genre, V., Kenny, G., Meyler, A. and Timmermann, A., 2013. Combining expert forecasts: Can anything beat the simple average?. International Journal of Forecasting, 29(1), pp.108-121.
[38] Giacomini, R., Skreta, V. and Turen, J., 2020. Heterogeneity, inattention, and Bayesian updates. American Economic Journal: Macroeconomics, 12(1), pp.282-309.
[39] Grossman, S.J. and Stiglitz, J.E., 1976. Information and competitive price systems. The American Economic Review, pp.246-253.
[40] Hellwig, M.F., 1980. On the aggregation of information in competitive markets. Journal of Economic Theory, 22(3), pp.477-498.
[41] Hellwig, C. and Veldkamp, L., 2009. Knowing what others know: Coordination motives in information acquisition. The Review of Economic Studies, 76(1), pp.223-251.
[42] Herbert, S., 2021. State-dependent central bank communication with heterogeneous beliefs. Available at SSRN 3923047.
[43] Kenny, G., Kostka, T. and Masera, F., 2014. How informative are the subjective density forecasts of macroeconomists?, Journal of Forecasting, 33(3), pp.163-185.
[44] Keane, M.P. and Runkle, D.E., 1990. Testing the rationality of price forecasts: New evidence from panel data. The American Economic Review, pp.714-735.
[45] Lahiri, K. and Sheng, X., 2010. Measuring forecast uncertainty by disagreement: The missing link. Journal of Applied Econometrics, 25(4), pp.514-538.
[46] Laster, D., Bennett, P. and Geoum, I.S., 1999. Rational bias in macroeconomic forecasts. The Quarterly Journal of Economics, 114(1), pp.293-318.
[47] Lorenzoni, G., 2009. A theory of demand shocks. American Economic Review, 99(5), pp.2050-2084.
[48] Lucas, R., 1972. Expectations and the neutrality of money. Journal of Economic Theory, 4(2), pp.103-124.
[49] Maćkowiak, B. and Wiederholt, M., 2009. Optimal sticky prices under rational inattention. American Economic Review, 99(3), pp.769-803.
[50] Maćkowiak, B. and Wiederholt, M., 2015. Business cycle dynamics under rational inattention. The Review of Economic Studies, 82(4), pp.1502-1532.
[51] Manzan, S., 2021. Are professional forecasters Bayesian?. Journal of Economic Dynamics and Control, 123, p.104045.
[52] Morris, S. and Shin, H.S., 2002. Social value of public information. American Economic Review, 92(5), pp.1521-1534.

[53] New York Times, 2012, Fed Ties Rates to Joblessness, With Target of 6.5%, https://www.nytimes.com/2012/12/13/business/economy/fed-to-maintain-stimulus-bond-buying.html.
[54] Nimark, K.P., 2014. Man-bites-dog business cycles. American Economic Review, 104(8), pp.2320-2367.
[55] Nimark, K.P. and Pitschner, S., 2019. News media and delegated information choice. Journal of Economic Theory, 181, pp.160-196.
[56] Ordoñez, G., 2013. The asymmetric effects of financial frictions. Journal of Political Economy, 121(5), pp.844-895.
[57] Pfäuti, O., 2023. The Inflation Attention Threshold and Inflation Surges. arXiv preprint arXiv:2308.09480.
[58] Rossi, B., Sekhposyan, T., and Soupre, M. (2019). "Understanding the Sources of Macroeconomic Uncertainty." Unpublished manuscript (updated version of Rossi et al. (2016) SSRN Working Paper No. 2816841).
[59] Shannon, C.E., 1948. A mathematical theory of communication. The Bell system technical journal, 27(3), pp.379-423.
[60] Sims, C.A., 1998, December. Stickiness. In Carnegie-Rochester conference series on public policy (Vol. 49, pp. 317-356). North-Holland.
[61] Sims, C.A., 2003. Implications of rational inattention. Journal of Monetary Economics, 50(3), pp.665-690.
[62] Song, W. and Stern, S., 2024. Firm inattention and the efficacy of monetary policy: A text-based approach. Review of Economic Studies, rdae102.
[63] Svensson, L.E.O., 2006. Social value of public information: Comment: Morris and shin (2002) is actually pro-transparency, not con. American Economic Review, 96(1), pp.448-452.
[64] Van Nieuwerburgh, S. and Veldkamp, L., 2006. Learning asymmetries in real business cycles. Journal of Monetary Economics, 53(4), pp.753-772.
[65] Veldkamp, L.L., 2005. Slow boom, sudden crash. Journal of Economic Theory, 2(124), pp.230-257.
[66] Wong, J., 2008. Information acquisition, dissemination, and transparency of monetary policy. Canadian Journal of Economics/Revue Canadienne d'Economique, 41(1), pp.46-79.
[67] Zarnowitz, V., 1979. An analysis of annual and multiperiod quarterly forecasts of aggregate income, output, and the price level. Journal of Business, pp.1-33.
[68] Zarnowitz, V., 1985. Rational expectations and macroeconomic forecasts. Journal of Business & Economic Statistics, 3(4), pp.293-311.
[69] Zarnowitz, V. and Braun, P., 1993. Twenty-two years of the NBER-ASA quarterly economic outlook surveys: aspects and comparisons of forecasting performance. In Business cycles, indicators, and forecasting (pp. 11-94). University of Chicago Press.
[70] Zarnowitz, V. and Lambros, L.A., 1987. Consensus and uncertainty in economic prediction. Journal of Political Economy, 95(3), pp.591-621.

Appendix A. Proofs

A.1. **Proof of Proposition 1.** We want to prove that the first order condition (4.1) is sufficient to identify the probabilities $p\left(\widehat{s}_t \mid \boldsymbol{x}\right) \in (0,1)^N$ that minimizes (3.4). The logic of the proof is as follows. Since the function to be minimized is smooth and defined on a closed set, its minimum will either be at the boundary or at an interior point where the first order condition (FOC) holds.. We will show that near the boundary, the function tends to infinity, so that the minimum must be at an interior point. We then show that at all interior points where the FOC holds, the function is convex. Hence, the FOC identifies a unique minimizing signal, up to a the normalizing constant..

To find the FOC, use that the log of a ratio is equal to the differences in logs, we can rewrite the minimization problem (3.4) as

$$p\left(\widehat{s}_t \mid \boldsymbol{x}\right) = \arg \min_{p(\widehat{s}_t \mid \boldsymbol{x}) \in (0,1)^N} \sum_{j=1}^{J} \sum_{n=1}^{N} p(x_n \mid \Omega_t^j) \left[\log\Big(\sum_{i=1}^{N} p\left(\widehat{s}_t \mid x_i\right) p(x_i \mid \Omega_{t-1}^j) \Big) - \log\Big(p\left(\widehat{s}_t \mid x_n\right) p(x_n \mid \Omega_{t-1}^j) \Big) \right] \tag{A.1}$$

The first order conditions w.r.t. $p\left(\widehat{s}_t \mid x_n\right)$ is then given by

$$\sum_{j=1}^{J} \sum_{m=1}^{N} p(x_m \mid \Omega_t^j) \frac{p(x_n \mid \Omega_{t-1}^j)}{\sum_{i=1}^{N} \Big(p(x_i \mid \Omega_{t-1}^j) p\left(\widehat{s}_t \mid x_i\right) \Big)} - \sum_{j=1}^{J} p(x_n \mid \Omega_t^j) \frac{p(x_n \mid \Omega_{t-1}^j)}{p(x_n \mid \Omega_{t-1}^j) p\left(\widehat{s}_t \mid x_n\right)} = 0 \tag{A.2}$$

which can be rearranged to the desired expression

$$\frac{1}{J} \sum_{j=1}^{J} p(x_n \mid \Omega_t^j) = \frac{1}{J} \sum_{j=1}^{J} \frac{p\left(\widehat{s}_t \mid x_n\right) p(x_n \mid \Omega_{t-1}^j)}{\sum_{i=1}^{N} \Big(p(x_i \mid \Omega_{t-1}^j) p\left(\widehat{s}_t \mid x_i\right) \Big)} \tag{A.3}$$

$$= \frac{1}{J} \sum_{j=1}^{J} p(x_n \mid \Omega_{t-1}^j, \widehat{s}_t). \tag{A.4}$$

From FOC we have the Jacobian of the objective function,

$$\nabla f = \left[\sum_{j=1}^{J} \frac{p(x_1 | \Omega_{t-1}^j)}{\sum_{i=1}^{N} \big(p(x_i | \Omega_{t-1}^j) p(\widehat{s}_t | x_i) \big)} - \frac{\sum_{j=1}^{J} p(x_1 | \Omega_t^j)}{p(\widehat{s}_t | x_1)} \quad \cdots \quad \sum_{j=1}^{J} \frac{p(x_N | \Omega_{t-1}^j)}{\sum_{i=1}^{N} \big(p(x_i | \Omega_{t-1}^j) p(\widehat{s}_t | x_i) \big)} - \frac{\sum_{j=1}^{J} p(x_N | \Omega_t^j)}{p(\widehat{s}_t | x_N)} \right]' \tag{A.5}$$

implying that element k, l of the Hessian $\mathbf{H}_f$ is given by

$$h_{f,k,l} = \sum_{j=1}^{J} \frac{-p(x_k \mid \Omega_{t-1}^j)^2}{\Big(\sum_{i=1}^{N} p(x_i \mid \Omega_{t-1}^j) p\left(\widehat{s}_t \mid x_i\right) \Big)^2} + \sum_{j=1}^{J} \frac{p(x_k \mid \Omega_t^j)}{p\left(\widehat{s}_t \mid x_k\right)^2} \text{ if } k = l \tag{A.6}$$

$$h_{f,k,l} = \sum_{j=1}^{J} \frac{-p(x_k \mid \Omega_{t-1}^j) p(x_k \mid \Omega_{t-1}^j)}{\Big(\sum_{i=1}^{N} p(x_i \mid \Omega_{t-1}^j) p\left(\widehat{s}_t \mid x_i\right) \Big)^2} \text{ if } k \neq l \tag{A.7}$$

If $\mathbf{H}_f$ was a positive definite matrix everywhere, the FOC would be both necessary and sufficient to characterize the minimum. However, while this is not the case, $\mathbf{H}_f$ is a positive semi-definite matrix at all points where first order conditions hold. To see that,

$$\lambda'\mathbf{H}_f\lambda = \sum_{j=1}^{J}\sum_{i=1}^{N} p(x_i \mid \Omega_t^j)\frac{\lambda_i}{p\left(\widehat{s}_t \mid x_i\right)^2} - \sum_{j=1}^{J}\frac{\left(\sum_{i=1}^{N} p(x_i \mid \Omega_{t-1}^j)\lambda_i\right)^2}{\left(\sum_{i=1}^{N} p(x_i \mid \Omega_{t-1}^j)p\left(\widehat{s}_t \mid x_i\right)\right)^2} \tag{A.8}$$

$$= \sum_{j=1}^{J}\sum_{i=1}^{N}\frac{p(x_i \mid \Omega_{t-1}^j)\lambda_i^2}{p\left(\widehat{s}_t \mid x_i\right)\left(\sum_{k=1}^{N} p(x_k \mid \Omega_{t-1}^j)p\left(\widehat{s}_t \mid x_k\right)\right)} - \sum_{j=1}^{J}\frac{\left(\sum_{i=1}^{N} p(x_i \mid \Omega_{t-1}^j)\lambda_i\right)^2}{\left(\sum_{i=1}^{N} p(x_i \mid \Omega_{t-1}^j)p\left(\widehat{s}_t \mid x_i\right)\right)^2} \tag{A.9}$$

$$= \sum_{j=1}^{J}\sum_{i=1}^{N}\frac{p(x_i \mid \Omega_{t-1}^j)^2\lambda_i^2}{p(x_i \mid \Omega_{t-1}^j)p\left(\widehat{s}_t \mid x_i\right)\left(\sum_{k=1}^{N} p(x_k \mid \Omega_{t-1}^j)p\left(\widehat{s}_t \mid x_k\right)\right)} - \sum_{j=1}^{J}\frac{\left(\sum_{i=1}^{N} p(x_i \mid \Omega_{t-1}^j)\lambda_i\right)^2}{\left(\sum_{i=1}^{N} p(x_i \mid \Omega_{t-1}^j)p\left(\widehat{s}_t \mid x_i\right)\right)^2} \tag{A.10}$$

$$\geq 0 \tag{A.11}$$

for any positive vector λ. The first equality come from simply multiplying through with λ. The second equality comes from substituting in the FOC (4.1). The inequality is implied by Sedrakyan's lemma, $\sum_{i=1}^{n}\frac{u_i^2}{v_i} \geq \frac{\left(\sum_{i=1}^{n} u_i\right)^2}{\sum_{i=1}^{n} v_i}$, with $u_i = p(x_i \mid \Omega_{t-1}^j)\lambda_i$, and $v_i = p(x_i \mid \Omega_{t-1}^j)p\left(\widehat{s}_t \mid x_i\right)\left(\sum_{k=1}^{N} p(x_k \mid \Omega_{t-1}^j)p\left(\widehat{s}_t \mid x_k\right)\right)$. The last line holds with equality only when λ_i is proportional to $p\left(\widehat{s}_t \mid x_i\right)\left(\sum_{k=1}^{N} p(x_k \mid \Omega_{t-1}^j)p\left(\widehat{s}_t \mid x_k\right)\right)$. The minimizing signal is thus only unique up to a normalization of the prior probability of observing the realized signal. Stated differently, there are many signals that obtains the minimum, but the ratios $p\left(\widehat{s}_t \mid x_i\right)/p\left(\widehat{s}_t \mid x_k\right)$ are uniquely determined. Since the KL-divergence between two distributions tend to infinity if either distribution tend to the degenerate one, the objective function tend to infinity near the boundary points of the simplex and is thus always greater at the boundary than at an interior critical point. The interior local minimum must then in fact also be the global minimum.

A.2. **Proof of Proposition 2.** We want to prove that if the prior beliefs of all forecasters coincide, the relative probability of observing $\widehat{s}_t$ in states n and m equals the mean-posterior-over-mean-prior odds ratio so that

$$\frac{p\left(\widehat{s}_t \mid x_n\right)}{p\left(\widehat{s}_t \mid x_m\right)} = R_m^n. \tag{A.12}$$

Start by taking the ratio of the posterior probabilities of state n and m induced by the signal

$$\frac{\frac{1}{J}\sum_{j=1}^{J} p(x_n \mid \Omega_t^j)}{\frac{1}{J}\sum_{j=1}^{J} p(x_m \mid \Omega_t^j)} = \frac{\frac{1}{J}\sum_{j=1}^{J}\frac{p(\widehat{s}_t|x_n)p(x_n|\Omega_{t-1}^j)}{\sum_{i=1}^{N}\left(p(\widehat{s}_t|x_i)p(x_i|\Omega_{t-1}^j)\right)}}{\frac{1}{J}\sum_{j=1}^{J}\frac{p(\widehat{s}_t|x_m)p(x_m|\Omega_{t-1}^j)}{\sum_{i=1}^{N}\left(p(\widehat{s}_t|x_i)p(x_i|\Omega_{t-1}^j)\right)}}. \tag{A.13}$$

Since by assumption, $p(\widehat{s} \mid \Omega_{t-1}^j) = p(\widehat{s} \mid \Omega_{t-1}^k)$ for all $j, k \in \{1, 2, ..., J\}$ this can be simplified to

$$\frac{\sum_{j=1}^J p(x_n \mid \Omega_t^j)}{\sum_{j=1}^J p(x_m \mid \Omega_t^j)} = \frac{\sum_{j=1}^J p\left(\widehat{s}_t \mid x_n\right) p(x_n \mid \Omega_{t-1}^j)}{\sum_{j=1}^J p\left(\widehat{s}_t \mid x_m\right) p(x_m \mid \Omega_{t-1}^j)}. \tag{A.14}$$

which after the rearranging yields the desired result since by definition

$$R_m^n = \left(\frac{\frac{1}{J} \sum_{j=1}^J p\left(x_n \mid \Omega_t^j\right)}{\frac{1}{J} \sum_{j=1}^J p\left(x_m \mid \Omega_t^j\right)} \right) / \left(\frac{\frac{1}{J} \sum_{j=1}^J p\left(x_n \mid \Omega_{t-1}^j\right)}{\frac{1}{J} \sum_{j=1}^J p\left(x_m \mid \Omega_{t-1}^j\right)} \right). \tag{A.15}$$

A.3. **Proof of Proposition 3.** For $j = 1, 2$, the FOC implies

$$\frac{p\left(\widehat{s}_t \mid x_n\right)}{p\left(\widehat{s}_t \mid x_m\right)} = \frac{p\left(x_n \mid \Omega_t^1\right) + p\left(x_n \mid \Omega_t^2\right)}{p\left(x_m \mid \Omega_t^1\right) + p\left(x_m \mid \Omega_t^2\right)} \times \frac{\frac{p\left(x_m \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^1\right)} + \frac{p\left(x_m \mid \Omega_{t-1}^2\right)}{p\left(s \mid \Omega_{t-1}^2\right)}}{\frac{p\left(x_n \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^1\right)} + \frac{p\left(x_n \mid \Omega_{t-1}^2\right)}{p\left(s \mid \Omega_{t-1}^2\right)}} \tag{A.16}$$

and the desired inequality

$$\frac{p\left(\widehat{s}_t \mid x_n\right)}{p\left(\widehat{s}_t \mid x_m\right)} > R_m^n \tag{A.17}$$

hence holds if

$$\frac{\frac{p\left(x_m \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^1\right)} + \frac{p\left(x_m \mid \Omega_{t-1}^2\right)}{p\left(s \mid \Omega_{t-1}^2\right)}}{\frac{p\left(x_n \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^1\right)} + \frac{p\left(x_n \mid \Omega_{t-1}^2\right)}{p\left(s \mid \Omega_{t-1}^2\right)}} > \frac{p\left(x_m \mid \Omega_{t-1}^1\right) + p\left(x_m \mid \Omega_{t-1}^2\right)}{p\left(x_n \mid \Omega_{t-1}^1\right) + p\left(x_n \mid \Omega_{t-1}^2\right)}. \tag{A.18}$$

To derive the conditions in the proposition, start by multiplying the term on the left hand side by $p\left(s \mid \Omega_{t-1}^1\right)$ to get

$$\frac{p\left(x_m \mid \Omega_{t-1}^1\right) + p\left(x_m \mid \Omega_{t-1}^2\right) \frac{p\left(s \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^2\right)}}{p\left(x_n \mid \Omega_{t-1}^1\right) + p\left(x_n \mid \Omega_{t-1}^2\right) \frac{p\left(s \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^2\right)}} > \frac{p\left(x_m \mid \Omega_{t-1}^1\right) + p\left(x_m \mid \Omega_{t-1}^2\right)}{p\left(x_n \mid \Omega_{t-1}^1\right) + p\left(x_n \mid \Omega_{t-1}^2\right)}. \tag{A.19}$$

Since this expression holds with equality if the ratio of prior probabilities of observing the signal $\frac{p\left(s \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^2\right)}$ equals 1, the desired inequality holds if the derivative of the left hand side of (4.6) with respect to $\frac{p\left(s \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^2\right)}$ is positive. Define the function f as

$$f\left(\frac{p\left(s \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^2\right)}\right) = \frac{p\left(x_m \mid \Omega_{t-1}^1\right) + p\left(x_m \mid \Omega_{t-1}^2\right) \frac{p\left(s \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^2\right)}}{p\left(x_n \mid \Omega_{t-1}^1\right) + p\left(x_n \mid \Omega_{t-1}^2\right) \frac{p\left(s \mid \Omega_{t-1}^1\right)}{p\left(s \mid \Omega_{t-1}^2\right)}}. \tag{A.20}$$

The quotient rule then gives

$$f' = \frac{p\left(x_m \mid \Omega^2_{t-1}\right)\left(p\left(x_n \mid \Omega^1_{t-1}\right) + p\left(x_n \mid \Omega^2_{t-1}\right)y\right) - p\left(x_n \mid \Omega^2_{t-1}\right)\left(p\left(x_m \mid \Omega^1_{t-1}\right) + p\left(x_m \mid \Omega^2_{t-1}\right)y\right)}{\left(p\left(x_n \mid \Omega^1_{t-1}\right) + p\left(x_n \mid \Omega^2_{t-1}\right)\frac{p\left(s \mid \Omega^1_{t-1}\right)}{p\left(s \mid \Omega^2_{t-1}\right)}\right)^2} \tag{A.21}$$

so that $f' > 0$ if

$$p\left(x_m \mid \Omega^2_{t-1}\right)\left(p\left(x_n \mid \Omega^1_{t-1}\right) + p\left(x_n \mid \Omega^2_{t-1}\right)\frac{p\left(s \mid \Omega^1_{t-1}\right)}{p\left(s \mid \Omega^2_{t-1}\right)}\right) -$$
$$p\left(x_n \mid \Omega^2_{t-1}\right)\left(p\left(x_m \mid \Omega^1_{t-1}\right) + p\left(x_m \mid \Omega^2_{t-1}\right)\frac{p\left(s \mid \Omega^1_{t-1}\right)}{p\left(s \mid \Omega^2_{t-1}\right)}\right) > 0 \tag{A.22}$$

which can be simplified to

$$\frac{p\left(x_m \mid \Omega^2_{t-1}\right)}{p\left(x_n \mid \Omega^2_{t-1}\right)} > \frac{p\left(x_m \mid \Omega^1_{t-1}\right)}{p\left(x_n \mid \Omega^1_{t-1}\right)}. \tag{A.23}$$

This implies that if we increase the ratio $\frac{p\left(s \mid \Omega^1_{t-1}\right)}{p\left(s \mid \Omega^2_{t-1}\right)}$ starting from 1, then the desired inequality follows if conditions (4.6) and (4.7) hold. (If only one of the conditions hold, the inequality switches direction.)

A.4. **Proof of Proposition 4.** From the first order condition (4.1) we know that

$$\frac{1}{J}\sum_{j=1}^{J} p(x_n \mid \widehat{s}_t, \Omega^j_{t-1}) = \frac{1}{J}\sum_{j=1}^{J} p(x_n \mid s^j_t, s_t, \Omega^j_{t-1}) \ \forall n. \tag{A.24}$$

i.e. the average beliefs conditional on the extracted common signal equals the average posteriors. We want to show that under the conditions in the proposition,

$$\lim_{J \xrightarrow{p} \infty} \frac{1}{J}\sum_{j=1}^{J} p(x_n \mid \widehat{s}_t, \Omega^j_{t-1}) = \lim_{J \xrightarrow{p} \infty} \frac{1}{J}\sum_{j=1}^{J} p(x_n \mid s_t, \Omega^j_{t-1}) \ \forall n. \tag{A.25}$$

That is, we want to derive conditions that ensure that $p\left(\widehat{s}_t \mid x_n\right)$ converges to $p\left(s_t \mid x_n\right)$ as $j \to \infty$.

Combining A.24 and A.25, it is therefore sufficient to show that the conditions in the proposition ensures that in the limit as $j \to \infty$

$$\lim_{j \xrightarrow{p} \infty} \frac{1}{J}\sum_{j=1}^{J} \frac{p(s^j_t \mid x_n)p(x_n \mid s_t, \Omega^j_{t-1})}{p(s^j_t \mid s_t, \Omega^j_{t-1})} = \lim_{J \xrightarrow{p} \infty} \frac{1}{J}\sum_{j=1}^{J} p(x_n \mid s_t, \Omega^j_{t-1}). \tag{A.26}$$

From the law of large numbers, it is equivalent to have

$$E\left[\frac{p(s^j_t \mid x_n)p(x_n \mid s_t, \Omega^j_{t-1})}{\sum_{n=1}^{N} p(s^j_t \mid x_n)p(x_n \mid s_t, \Omega^j_{t-1})}\right] = E\left[p(x_n \mid s_t, \Omega^j_{t-1})\right]. \tag{A.27}$$

Under the assumed conditions,

$$E\left[\frac{p(s_t^j \mid x_n)p(x_n \mid s_t, \Omega_{t-1}^j)}{\sum_{n=1}^N p(s_t^j \mid x_n)p(x_n \mid s_t, \Omega_{t-1}^j)}\right] = E\left[\frac{p(s_t^j \mid x_n)}{\sum_{n=1}^N p(s_t^j \mid x_n)}\right] \tag{A.28}$$

$$= \frac{1}{N}. \tag{A.29}$$

The first equality comes from plugging in$p(x_n \mid s_t, \Omega_{t-1}^j) = \frac{1}{N}$. The second equality holds because $\frac{p(s_t^j|x_n)}{\sum_{n=1}^N p(s_t^j|x_n)}$are identically distributed across n and sum to 1.

A.5. **Proof of Proposition 6.** The strategy of the proof is to first derive the cross-sectional average posterior distribution for the linear Gaussian information structure. We then use that the first order condition from the KL-minimization problem states that this should equal the cross-sectional average posterior implied by the estimated common signal $\hat{s}$. This reduced the problem to matching coefficients across distributions.

The linear-Gaussian filtering problem described by (4.10) - (4.11) implies that agent j's posterior is given by $p\left(x \mid s, s^j, \underline{\mu}^j\right) = N\left(\bar{\mu}^j, \bar{\sigma}^2\right)$ where

$$\bar{\mu}^j = \frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}}\underline{\mu}^j + \frac{\sigma_\eta^{-2}}{\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}}s + \frac{\sigma_\varepsilon^{-2}}{\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}}s^j \tag{A.30}$$

and

$$\bar{\sigma}^2 = \left(\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}\right)^{-1}. \tag{A.31}$$

The cross-sectional average distribution is the integral of the compound distribution of two normal distributions, which again is a normal distribution and given by

$$\int_j p\left(x \mid s, s^j, \underline{\mu}^j\right) p\left(s^j, \underline{\mu}^j\right) dj = N\left(g_\mu \underline{\mu} + g_s s + g_j x,\ g_\mu^2 \sigma_\mu^2 + g_j^2 \sigma_\varepsilon^2 + \left(\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}\right)^{-1}\right) \tag{A.32}$$

where

$$g_\mu = \frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}}, g_s = \frac{\sigma_\eta^{-2}}{\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}}, g_j = \frac{\sigma_\varepsilon^{-2}}{\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}}. \tag{A.33}$$

To prove the proposition, we need to find a signal $\widehat{s}$ with conditional distribution $p\left(\widehat{s} \mid \boldsymbol{x}\right)$ such that the implied average posterior is equal to (A.32) above. If

$$\widehat{s} = x + \widehat{\eta} : \widehat{\eta} \sim N\left(0, \widehat{\sigma}_\eta^{-2}\right) \tag{A.34}$$

then the posterior belief of agent j is given by

$$p\left(x \mid \widehat{s}, \underline{\mu}^j\right) = N\left(\frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2}}\underline{\mu}^j + \frac{\widehat{\sigma}_\eta^{-2}}{\underline{\sigma}^{-2} + \hat{\sigma}_\eta^{-2}}\widehat{s}, \left(\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2}\right)^{-1}\right) \tag{A.35}$$

implying an average and that

$$\underline{\mu}^j = \underline{\mu} + \underline{\varepsilon}^j : \underline{\varepsilon}^j \sim N\left(0, \sigma_\mu^2\right) \tag{A.36}$$

so that we can write

$$x = g\underline{\mu} + (1-g)\,\widehat{s} + g\underline{\varepsilon}^j + \delta : \delta \sim N\left(0, \left(\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2}\right)^{-1}\right) \tag{A.37}$$

so that

$$x \mid \widehat{s} \sim N\left(\widehat{g}\underline{\mu} + (1-\widehat{g})\,\widehat{s}, \widehat{g}^2\sigma_\mu^2 + \left(\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2}\right)^{-1}\right) \tag{A.38}$$

$$= \int_j p\left(x \mid \widehat{s}, \underline{\mu}^j\right) dj \tag{A.39}$$

where

$$\widehat{g} = \frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2}}. \tag{A.40}$$

Equating the two cross-sectional average posteriors

$$\int_j p\left(x \mid s, s^j, \underline{\mu}^j\right) p\left(s^j, \underline{\mu}^j\right) dj = \int_j p\left(x \mid \widehat{s}, \underline{\mu}^j\right) dj \tag{A.41}$$

we get a system of two equations and two unknowns

$$g_\mu\underline{\mu} + g_s s + g_j x = \widehat{g}\underline{\mu} + (1-\widehat{g})\,\widehat{s} \tag{A.42}$$

$$g_\mu^2\sigma_\mu^2 + g_j^2\sigma_\varepsilon^2 + \left(\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}\right)^{-1} = \widehat{g}^2\sigma_\mu^2 + \left(\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2}\right)^{-1} \tag{A.43}$$

Rearranging these expressions gives the desired result.

Alternative derivation. If

$$\mu^j \sim N(\mu, \sigma_\mu^2), \quad \mu \text{ is a constant,} \tag{A.44}$$

$$x \mid s, \mu^j \sim N\left(a\mu^j + (1-a)s, \sigma_x^2\right) \tag{A.45}$$

then, $x \mid s \sim N\left(\widehat{g}\mu + (1-\widehat{g})s, \sigma_x^2 + \widehat{g}^2\sigma_\mu^2\right)$.

$$p(x \mid s, \mu^j) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{\frac{\left(x - \widehat{g}\mu^j - (1-\widehat{g})s\right)^2}{2\sigma_x^2}}, \tag{A.46}$$

$$p(x \mid s) = \int_{\mathbb{R}} p(x \mid s, \mu^j) p(\mu^j \mid s) d\mu^j = \int_{\mathbb{R}} p(x \mid s, \mu^j) p(\mu^j) d\mu^j \tag{A.47}$$

because the prior distribution does not depend on s.

The following is an alternative derivation to the one above, which yields the same end result. The benefit of this second derivation is that it shows explicitly why the compound distribution consisting of a continuum of normal distributions with normal distributed means is also a normal distribution.

Start by expanding the integral (A.47) to get

$$p(x \mid s) = \int \frac{1}{\sqrt{2\pi}\sigma_x} e^{\frac{\left(x - \widehat{g}\mu^j - (1-\widehat{g})s\right)^2}{2\sigma_x^2}} \frac{1}{\sqrt{2\pi}\sigma_\mu} e^{\frac{\left(\mu^j - \mu\right)^2}{2\sigma_\mu^2}} d\mu^j \tag{A.48}$$

and combine the two product of the two exponential terms

$$p(x \mid s) = \int \frac{1}{2\pi\sigma_x\sigma_\mu} e^{\frac{\sigma_\mu^2\left(x - \widehat{g}\mu^j - (1-\widehat{g})s\right)^2 + \sigma_x^2\left(\mu^j - \mu\right)^2}{2\sigma_x^2\sigma_\mu^2}} d\mu^j. \tag{A.49}$$

Expand the squared terms

$$p(x \mid s)=\int \frac{1}{2\pi\sigma_x\sigma_\mu}e^{\frac{(\widehat{g}^2\sigma_\mu^2+\sigma_x^2){\mu^j}^2-(2\widehat{g}x\sigma_\mu^2-2\widehat{g}(1-\widehat{g})s\sigma_\mu^2+2\mu\sigma_x^2)\mu^j+(\sigma_\mu^2x^2+\sigma_\mu^2(1-\widehat{g})^2s^2-2(1-\widehat{g})sx\sigma_\mu^2+\sigma_x^2\mu^2)}{2\sigma_x^2\sigma_\mu^2}}d\mu^j$$

and isolate the constants and divide by $2\sigma_x^2\sigma_\mu^2$ so that

$$p(x \mid s)=e^{\left(\frac{(\sigma_\mu^2x^2+\sigma_\mu^2(1-\widehat{g})^2s^2-2(1-\widehat{g})sx\sigma_\mu^2+\sigma_x^2\mu^2)}{2\sigma_x^2\sigma_\mu^2}\right)}\int \frac{1}{2\pi\sigma_x\sigma_\mu}e^{\left(\frac{\left(\mu^j\right)^2-\frac{(2\widehat{g}x\sigma_\mu^2-2\widehat{g}(1-\widehat{g})s\sigma_\mu^2+2\mu\sigma_x^2)}{(a^2\sigma_\mu^2+\sigma_x^2)}\mu^j}{\frac{2\sigma_x^2\sigma_\mu^2}{(a^2\sigma_\mu^2+\sigma_x^2)}}\right)}d\mu^j$$

This can be rearranged to be a Gaussian density for μ^j

$$p(x \mid s)=\frac{1}{\sqrt{2\pi}\sqrt{\widehat{g}^2\sigma_\mu^2+\sigma_x^2}}e^{\left(\frac{(\sigma_\mu^2x^2+\sigma_\mu^2(1-\widehat{g})^2s^2-2(1-\widehat{g})sx\sigma_\mu^2+\sigma_x^2\mu^2)}{2\sigma_x^2\sigma_\mu^2}-\frac{(\widehat{g}x\sigma_\mu^2-\widehat{g}(1-\widehat{g})s\sigma_\mu^2+\mu\sigma_x^2)^2}{2\sigma_x^2\sigma_\mu^2(\widehat{g}^2\sigma_\mu^2+\sigma_x^2)}\right)} \tag{A.50}$$

$$\times\int \frac{1}{\sqrt{2\pi}}\frac{\sqrt{\widehat{g}^2\sigma_\mu^2+\sigma_x^2}}{\sigma_x\sigma_\mu}e^{\left(\frac{\left(\mu^j-\frac{(\widehat{g}x\sigma_\mu^2-\widehat{g}(1-\widehat{g})s\sigma_\mu^2+\mu\sigma_x^2)}{(\widehat{g}^2\sigma_\mu^2+\sigma_x^2)}\right)^2}{\frac{2\sigma_x^2\sigma_\mu^2}{(\widehat{g}^2\sigma_\mu^2+\sigma_x^2)}}\right)}d\mu^j \tag{A.51}$$

i.e. the previous expression is of the form of the form $\frac{1}{\sqrt{2\pi}\sigma}\exp\frac{\left(\mu^j-\mu\right)^2}{2\sigma^2}$. To simplify, merge the constants

$$p(x \mid s)=\frac{1}{\sqrt{2\pi}\sqrt{\widehat{g}^2\sigma_\mu^2+\sigma_x^2}}e^{\left(\frac{(\sigma_\mu^2x^2+\sigma_\mu^2(1-\widehat{g})^2s^2-2(1-\widehat{g})sx\sigma_\mu^2+\sigma_x^2\mu^2)}{2\sigma_x^2\sigma_\mu^2}-\frac{(\widehat{g}x\sigma_\mu^2-\widehat{g}(1-\widehat{g})s\sigma_\mu^2+\mu\sigma_x^2)^2}{2\sigma_x^2\sigma_\mu^2(\widehat{g}^2\sigma_\mu^2+\sigma_x^2)}\right)} \tag{A.52}$$

Combine the terms in the exponential

$$p(x \mid s)=\frac{1}{\sqrt{2\pi}\sqrt{\widehat{g}^2\sigma_\mu^2+\sigma_x^2}}e^{\left(\frac{x^2-2((1-\widehat{g})s+\widehat{g}\mu)+\widehat{g}^2\mu^2+(1-\widehat{g})^2s^2-2\widehat{g}(1-\widehat{g})s\mu}{2(\widehat{g}^2\sigma_\mu^2+\sigma_x^2)}\right)} \tag{A.53}$$

and simplify to get

$$p(x \mid s)=\frac{1}{\sqrt{2\pi}\sqrt{\widehat{g}^2\sigma_\mu^2+\sigma_x^2}}e^{\left(\frac{(x-((1-\widehat{g})s+\widehat{g}\mu))^2}{2(\widehat{g}^2\sigma_\mu^2+\sigma_x^2)}\right)} \tag{A.54}$$

What we have in (A.54) is then the pdf of a normal distribution with mean $(1-\widehat{g})s+\widehat{g}\mu$ and variance $\widehat{g}^2\sigma_\mu^2+\sigma_x^2$.

A.6. **Proof of Proposition 5.** For each agent j, combining the prior $x|\Omega_{t-1}^j\sim\mathcal{N}(\underline{\mu}^j,\underline{\sigma}^2)$ with signals (s_t,s_t^j) via Gaussian Bayesian-updating rules yields a posterior mean

$$E\left(x \mid \Omega_{t-1}^j,s_t,s_t^j\right)=\frac{\frac{1}{\underline{\sigma}^2}\underline{\mu}^j+\frac{1}{\sigma_\eta^2}s_t+\frac{1}{\sigma_\varepsilon^2}s_t^j}{\frac{1}{\underline{\sigma}^2}+\frac{1}{\sigma_\eta^2}+\frac{1}{\sigma_\varepsilon^2}}. \tag{A.55}$$

Observe that this equation depends linearly on s_t and s_t^j.

We have J observed posterior means $E\left(x \mid \Omega_{t-1}^{1}, s_t, s_t^1\right), \ldots, E\left(x \mid \Omega_{t-1}^{J}, s_t, s_t^J\right)$, but there are $J+1$ unobserved signals $\{s_t, s_t^1, \ldots, s_t^J\}$. Hence there are infinitely many solutions in general. Fix one solution

$$\left(\hat{s}_t, \hat{s}_t^1, \ldots, \hat{s}_t^J\right) \tag{A.56}$$

that reproduces the observed $\{E\left(x \mid \Omega_{t-1}^{j}, s_t, s_t^j\right)\}$. Consider another candidate solution

$$\left(\hat{s}_t + c, \hat{s}_t^1 - \alpha_1 c, \ldots, \hat{s}_t^J - \alpha_J c\right) \tag{A.57}$$

for constants $c \neq 0$and $\{\alpha_j\}$ chosen so the linear combination (posterior mean) remains unchanged. One can solve for $\{\alpha_j\}$ to ensure each $E\left(x \mid \Omega_{t-1}^{j}, s_t, s_t^j\right)$ stays fixed.

Because there is a whole family of $\left(s_t, s_t^j\right)$ combinations mapping to the same$E\left(x \mid \Omega_{t-1}^{j}, s_t, s_t^j\right)$ values, the realized common signal s_t is not uniquely identified by the observed cross-section of posterior means.

A.7. **Proof of Corollary 2.** Take the ratio of the averages of (4.8) across agents for state n and m, drop the indices for the prior information set,

$$\frac{\sum_{j=1}^{J} p_j(x_n \mid \Omega_{t-1}, s_t)}{\sum_{j=1}^{J} p_j(x_m \mid \Omega_{t-1}, s_t)} = \frac{\sum_{j=1}^{J}(p_j(s_t \mid x_n)p(x_n \mid \Omega_{t-1}))}{\sum_{j=1}^{J}(p_j(s_t \mid x_m)p(x_m \mid \Omega_{t-1}))} \tag{A.58}$$

$$= \frac{\sum_{j=1}^{J} p_j(s_t \mid x_n)}{\sum_{j=1}^{J} p_j(s_t \mid x_m)} \frac{p(x_n \mid \Omega_{t-1})}{p(x_m \mid \Omega_{t-1})} \tag{A.59}$$

Rewrite the equality to get

$$\frac{\sum_{j=1}^{J} p_j(s_t \mid x_n)}{\sum_{j=1}^{J} p_j(s_t \mid x_m)} = \frac{\frac{\sum_{j=1}^{J} p_j(x_n|\Omega_{t-1}, s_t)}{p(x_n|\Omega_{t-1})}}{\frac{\sum_{j=1}^{J} p_j(x_m|\Omega_{t-1}, s_t)}{p(x_m|\Omega_{t-1})}} \tag{A.60}$$

From (A.3) in the proof of Proposition 1, we can write

$$\frac{p\left(\widehat{s}_t \mid x_n\right)}{p\left(\widehat{s}_t \mid x_m\right)} = \frac{\sum_{j=1}^{J} p\left(x_n \mid \Omega_t^j\right)}{\sum_{j=1}^{J} p\left(x_m \mid \Omega_t^j\right)} \frac{\sum_{j=1}^{J} \frac{p\left(x_m|\Omega_{t-1}^j\right)}{\sum_{i=1}^{N}\left(p\left(x_i|\Omega_{t-1}^j\right)p(\widehat{s}_t|x_i)\right)}}{\sum_{j=1}^{J} \frac{p\left(x_n|\Omega_{t-1}^j\right)}{\sum_{i=1}^{N}\left(p\left(x_i|\Omega_{t-1}^j\right)p(\widehat{s}_t|x_i)\right)}} \tag{A.61}$$

$$= \frac{\frac{\sum_{j=1}^{J} p\left(x_n|\Omega_t^j\right)}{p(x_n|\Omega_{t-1})}}{\frac{\sum_{j=1}^{J} p\left(x_m|\Omega_t^j\right)}{p(x_m|\Omega_{t-1})}} \tag{A.62}$$

That gives us the desired equality because in the heterogeneous likelihood case, the factual posterior $p\left(x_n \mid \Omega_t^j\right)$ equals $p_j(x_n \mid \Omega_{t-1}, s_t)$.

A.8. **Proof of Corollary 3.** We want to show that when $\sigma_\varepsilon^2 \to \infty, \widehat{s} \to s$ and $\widehat{\sigma}_\eta^2 \to \sigma_\eta^2$. From (4.14) we then have that

$$g_\mu \to \frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2}+\sigma_\eta^{-2}}, g_s \to \frac{\sigma_\eta^{-2}}{\underline{\sigma}^{-2}+\sigma_\eta^{-2}}, g_j \to 0, \tag{A.63}$$

as $\sigma_\varepsilon^2 \to \infty$. Plug in (4.17)

$$(\frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2}+\sigma_\eta^{-2}})^2 \cdot \sigma_\mu^2 + (\underline{\sigma}^{-2}+\sigma_\eta^{-2})^{-1} = (\frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2}+\widehat{\sigma}_\eta^{-2}})^2 \cdot \sigma_\mu^2 + (\underline{\sigma}^{-2}+\widehat{\sigma}_\eta^{-2})^{-1} \tag{A.64}$$

Notice that $\widehat{\sigma}_\eta^2 = \sigma_\eta^2$ is one solution to this equation. It is also the unique solution because the right hand side of (A.64)is increasing in $\widehat{\sigma}_\eta^2$, while the left hand side is fixed.
Thus, $\widehat{g} \to \frac{\underline{\sigma}^{-2}}{\sigma_\eta^{-2}+\underline{\sigma}^{-2}}$, and $\widehat{s} = (1-\widehat{g})^{-1}\left[(g_\mu - \widehat{g})\,\underline{\mu} + g_s s + g_j x\right] \to s$.
Now we show the other direction. From 4.16 we can see that $\widehat{s} = s$ for all realizations of s,if and only if $(1-\widehat{g})^{-1} g_s = 1$, and $(1-\widehat{g})^{-1}\left[(g_\mu - \widehat{g})\,\underline{\mu} + g_j x\right] = 0$.
$(1-\widehat{g})^{-1} g_s = 1$ gives

$$\widehat{\sigma}_\eta^{-2} = \frac{\sigma_\eta^{-2}}{\sigma_\varepsilon^{-2}+\sigma_\eta^{-2}+\underline{\sigma}^{-2}}(\sigma_\eta^{-2}+\underline{\sigma}^{-2})$$

Since the mean of the prior dispersion $\underline{\mu}$ and true state x are not restricted, we can infer from $(1-\widehat{g})^{-1}\left[(g_\mu - \widehat{g})\,\underline{\mu} + g_j x\right] = 0$ that $g_\mu = \widehat{g}$ and $g_j = 0$. Combining all the conditions gives us $\sigma_\varepsilon^2 = \infty$.

A.9. **Proof of Corollary 4.** If the true common signal is uninformative, i.e., $\sigma_\eta^2 \to \infty$,then

$$g_\mu \to \frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2}+\sigma_\varepsilon^{-2}}, g_s \to 0, g_j \to \frac{\sigma_\varepsilon^{-2}}{\underline{\sigma}^{-2}+\sigma_\varepsilon^{-2}}, \tag{A.65}$$

Plug in (4.17)

$$(\frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2}+\sigma_\varepsilon^{-2}})^2 \cdot \sigma_\mu^2 + (\frac{\sigma_\varepsilon^{-2}}{\underline{\sigma}^{-2}+\sigma_\varepsilon^{-2}})^2 \cdot \sigma_\varepsilon^2 + (\underline{\sigma}^{-2}+\sigma_\varepsilon^{-2})^{-1} = (\frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2}+\hat{\sigma}_\eta^{-2}})^2 \cdot \sigma_\mu^2 + (\underline{\sigma}^{-2}+\hat{\sigma}_\eta^{-2})^{-1} \tag{A.66}$$

Again this equation has a unique solution for $\widehat{\sigma}_\eta^2$, note also this solution must be $\widehat{\sigma}_\eta^2 > \sigma_\varepsilon^2$ by monotonicity of the right-hand side of (A.66) in $\widehat{\sigma}_\eta^2$. Therefore, $g_\mu = \frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2}+\sigma_\varepsilon^{-2}} < \frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2}+\widehat{\sigma}_\eta^{-2}} = \widehat{g}$.
From (4.16) we then have the desired result

$$\begin{aligned} \widehat{s} &= (1-\widehat{g})^{-1}\left[(g_\mu - \widehat{g})\,\underline{\mu} + g_s s + g_j x\right] && \text{(A.67)} \\ &= (1-\widehat{g})^{-1} g_j \left[\frac{g_\mu - \widehat{g}}{g_j}\underline{\mu} + x\right] && \text{(A.68)} \\ &= \alpha(x - \beta\underline{\mu}) && \text{(A.69)} \end{aligned}$$

where $\alpha = (1-\widehat{g})^{-1} g_j$, $\beta = \frac{\widehat{g}-g_\mu}{g_j}$, and $\alpha > (1-g_\mu)^{-1} g_j = 1$, $\beta < \frac{1-g_\mu}{g_j} = 1$, $\alpha\beta < 1$.

A.10. **Proof of Corollary 5.** To show the estimated precision $\widehat{\sigma}_\eta^{-2}$ is increasing in σ_ε^{-2} and σ_η^{-2}, we apply the Implicit Function Theorem to 4.17.
We start by showing that $\partial\widehat{\sigma}_\eta^{-2}/\partial\sigma_\eta^{-2} > 0$. The derivative of the right-hand side of 4.17with respect to the precision of the estimated common signal $\widehat{\sigma}_\eta^{-2}$ is

$$\frac{\partial RHS}{\partial\widehat{\sigma}_\eta^{-2}} = \frac{\partial\left((\frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2}+\widehat{\sigma}_\eta^{-2}})^2\cdot\sigma_\mu^2 + (\underline{\sigma}^{-2}+\widehat{\sigma}_\eta^{-2})^{-1}\right)}{\partial\widehat{\sigma}_\eta^{-2}} \tag{A.70}$$

$$= \underbrace{\frac{\partial\left((\frac{\underline{\sigma}^{-2}}{\underline{\sigma}^{-2}+\widehat{\sigma}_\eta^{-2}})^2\cdot\sigma_\mu^2\right)}{\partial\widehat{\sigma}_\eta^{-2}}}_{<0} + \underbrace{\frac{\partial\left((\underline{\sigma}^{-2}+\widehat{\sigma}_\eta^{-2})^{-1}\right)}{\partial\widehat{\sigma}_\eta^{-2}}}_{<0} \tag{A.71}$$

$$< 0 \tag{A.72}$$

and the derivative of the left-hand side w.r.t. the precision of the true common signal σ_η^{-2} is

$$\frac{\partial LHS}{\partial\sigma_\eta^{-2}} = \frac{\partial\left(g_\mu^2\sigma_\mu^2 + g_j^2\sigma_\varepsilon^2 + \left(\underline{\sigma}^{-2}+\sigma_\eta^{-2}+\sigma_\varepsilon^{-2}\right)^{-1}\right)}{\partial\sigma_\eta^{-2}} \tag{A.73}$$

$$= \underbrace{\frac{\partial g_\mu^2\sigma_\mu^2}{\partial\sigma_\eta^{-2}}}_{<0} + \underbrace{\frac{\partial g_j^2\sigma_\varepsilon^2}{\partial\sigma_\eta^{-2}}}_{<0} + \underbrace{\frac{\partial\left(\underline{\sigma}^{-2}+\sigma_\eta^{-2}+\sigma_\varepsilon^{-2}\right)^{-1}}{\partial\sigma_\eta^{-2}}}_{<0} \tag{A.74}$$

$$< 0 \tag{A.75}$$

To show the estimated precision $\widehat{\sigma}_\eta^{-2}$ is increasing in σ_ε^{-2} , we again apply the Implicit Function Theorem to 4.17, but with more arduous algebra.
The derivative of the right-hand side is the same as above. For the left-hand side, the derivative with respect to the precision of the private signal σ_ε^{-2} is

$$\frac{\partial LHS}{\partial\sigma_\varepsilon^{-2}} = \frac{\partial\left(g_\mu^2\sigma_\mu^2 + g_j^2\sigma_\varepsilon^2 + \left(\underline{\sigma}^{-2}+\sigma_\eta^{-2}+\sigma_\varepsilon^{-2}\right)^{-1}\right)}{\partial\sigma_\varepsilon^{-2}} \tag{A.76}$$

$$= \underbrace{\frac{\partial g_\mu^2\sigma_\mu^2}{\partial\sigma_\varepsilon^{-2}}}_{<0} + \frac{\partial\left(g_j^2\sigma_\varepsilon^2 + \left(\underline{\sigma}^{-2}+\sigma_\eta^{-2}+\sigma_\varepsilon^{-2}\right)^{-1}\right)}{\partial\sigma_\varepsilon^{-2}} \tag{A.77}$$

$$= \underbrace{\frac{\partial g_\mu^2\sigma_\mu^2}{\partial\sigma_\varepsilon^{-2}}}_{<0} + \frac{\partial\left[\left(\underline{\sigma}^{-2}+\sigma_\eta^{-2}+2\sigma_\varepsilon^{-2}\right)\left(\underline{\sigma}^{-2}+\sigma_\eta^{-2}+\sigma_\varepsilon^{-2}\right)^{-2}\right]}{\partial\sigma_\varepsilon^{-2}} \tag{A.78}$$

$$= \underbrace{\frac{\partial g_\mu^2\sigma_\mu^2}{\partial\sigma_\varepsilon^{-2}}}_{<0} + \underbrace{\frac{-2\sigma_\varepsilon^{-2}}{\left(\underline{\sigma}^{-2}+\sigma_\eta^{-2}+\sigma_\varepsilon^{-2}\right)^3}}_{<0} \tag{A.79}$$

$$< 0 \tag{A.80}$$

A.11. **Proof of Corollary 6.** We want to find the estimated private signal follow a conditional Gaussian distribution

$$\widehat{s^j} \mid x \sim N\left(x, \widehat{\sigma}_\varepsilon^2\right) \tag{A.81}$$

Intuitively, the estimated private signal should fill the gap between the true posterior and the posterior implied by the estimated common signal.
The true individual posterior follows a Gaussian distribution

$$N\left(g_\mu \underline{\mu}^j + g_s s + g_j s^j, \left(\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}\right)^{-1}\right) \tag{A.82}$$

On the other hand, based on (A.35) the individual posterior implied only by the estimated common signal is

$$N\left(\widehat{g}\underline{\mu}^j + (g_\mu - \widehat{g})\underline{\mu} + g_s s + g_j s^j, \left(\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2}\right)^{-1}\right) \tag{A.83}$$

Therefore, to match the variance

$$\widehat{\sigma}_\varepsilon^{-2} = \left(\underline{\sigma}^{-2} + \sigma_\eta^{-2} + \sigma_\varepsilon^{-2}\right) - \left(\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2}\right) \tag{A.84}$$

$$= \left(\sigma_\eta^{-2} - \widehat{\sigma}_\eta^{-2}\right) + \sigma_\varepsilon^{-2} \tag{A.85}$$

To match the mean, we apply the Bayesian updating formula

$$g_\mu \underline{\mu}^j + g_s s + g_j s^j = \frac{\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2}}{\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2} + \widehat{\sigma}_\varepsilon^{-2}} \cdot \left(\widehat{g}\underline{\mu}^j + (g_\mu - \widehat{g})\underline{\mu} + g_s s + g_j s^j\right) + \frac{\widehat{\sigma}_\varepsilon^{-2}}{\underline{\sigma}^{-2} + \widehat{\sigma}_\eta^{-2} + \widehat{\sigma}_\varepsilon^{-2}} \cdot \widehat{s^j} \tag{A.86}$$

Plug in $\widehat{\sigma}_\varepsilon^{-2}$and reorganize the terms, we have

$$\widehat{s^j} = g_\mu \underline{\mu} + g_s s + g_j s^j \tag{A.87}$$

The variance in (4.18) is simply given by equating the posterior variance implied by s and s^j with the posterior variance implied by $\widehat{s}$ and $\widehat{s^j}$.

Appendix B. Additional Figures

Figure B.1. Time series of informativeness of individual and common signals about PCE inflation, GDP deflator and complete sample for GDP growth.